

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OW protein - nucleic search, using frame_plus_p2n model

Run on: March 4, 2005, 22:43:16 ; Search time 230 Seconds
(without alignments)
3251.211 Million cell updates/sec

Title: US-10-691-374-2

Perfect score: 2390

Sequence: 1 MWIGNSILLIIRASSIS.....VTPHCTSLSEIRDEALVNY 457

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405566

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlp
-Q/cgn2_1/USPTO.spool.p/US10691374/runatc.28022005.120706.2101:/app_query.fasta_1.647
-DB=Issued Patents NA -QFMT=fastap -SUPFIX=eni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=Bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCNALGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10691374@cgn2_1.1.105@runatc.28022005.120706.2101 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:*
1: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
4: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
5: /cgn2_6/prodata/1/ina/FACTUS_COMB.seq:*
6: /cgn2_6/prodata/1/ina/backfillseq.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2390	100.0	1636	6	5447867-2 Patent No. 5447867
2	2390	100.0	1636	6	5447867-2 Patent No. 5447867
3	1113	46.6	1631	3	US-09-051-239A-1 Sequence 1, App11
4	1113	46.6	1631	4	US-10-151-668-1 Sequence 1, App11
5	1103	46.2	1657	3	US-08-941-532-5 Sequence 5, App11
6	787	32.9	1726	3	US-08-467-023-133 Sequence 133, App
7	786.5	32.9	1479	3	US-08-467-023-141 Sequence 141, App
8	768	32.1	1395	3	US-08-467-023-140 Sequence 140, App
9	768	32.1	1410	3	US-08-467-023-139 Sequence 139, App
10	366	15.3	219	1	US-08-463-213-1 Patent No. 5453566
11	366	15.3	219	6	5453566-1 Patent No. 5453566
12	366	15.3	219	6	5453566-1 Patent No. 5453566

13	299	12.5	1356	4	US-09-107-532A-2998 Sequence 2998, App
14	296	12.4	1086	4	US-09-614-221A-573 Sequence 573, App
15	293.5	12.3	1570	4	US-09-787-583-1 Sequence 1, App11
16	241	10.1	1611	1	US-08-061-062A-5 Sequence 5, App11
17	241	10.1	1611	3	US-08-536-150-5 Sequence 5, App11
18	221	9.2	2974	1	US-08-290-978A-4 Sequence 4, App11
19	221	9.2	2974	2	US-08-780-869-4 Sequence 4, App11
20	212.5	8.9	1329	4	US-09-107-532A-2600 Sequence 2600, App
21	205.5	8.6	1248	3	US-09-198-955-5 Sequence 5, App11
22	205.5	8.6	1248	3	US-09-670-141-5 Sequence 5, App11
23	192	8.0	864	3	US-09-087-194-34 Sequence 34, App1
24	189	7.9	710	3	US-08-941-532-7 Sequence 7, App11
25	186	7.8	710	3	US-08-998-416-603 Sequence 603, App
26	169	7.1	3250	1	US-08-061-062A-7 Sequence 7, App11
27	169	7.1	3250	1	US-08-536-150-7 Sequence 7, App11
28	164	6.9	219	3	US-09-051-239A-9 Sequence 9, App11
29	164	6.9	219	3	US-10-151-668-9 Sequence 9, App11
30	146.5	6.1	1664976	4	US-08-916-421B-1 Sequence 1, App11
31	146.5	6.1	1664976	4	US-09-692-570-1 Sequence 1, App11
32	134	5.6	155	3	US-09-051-239A-7 Sequence 7, App11
33	134	5.6	155	4	US-10-151-668-7 Sequence 7, App11
34	129.5	5.4	3018	3	US-09-206-942-40 Sequence 40, App1
35	129.5	5.4	3036	3	US-09-206-942-38 Sequence 38, App1
36	127	5.3	2013	4	US-09-107-433-2373 Sequence 2373, App
37	127	5.3	2127	4	US-09-583-110-358 Sequence 358, App
38	126	5.3	2577	4	US-09-733-643B-1 Sequence 1, App11
39	124.5	5.2	2706	4	US-09-134-001C-2514 Sequence 2514, App
40	123.5	5.2	1779	4	US-09-302-626B-5 Sequence 5, App11
41	121	5.1	1483	3	US-08-961-527-315 Sequence 315, App
42	120.5	5.0	3015	3	US-09-206-942-56 Sequence 56, App1
43	120.5	5.0	3033	3	US-09-206-942-54 Sequence 54, App1
44	118.5	5.0	3222	3	US-09-206-942-48 Sequence 48, App1
45	118.5	5.0	3240	3	US-09-206-942-46 Sequence 46, App1

ALIGNMENTS

RESULT 1
5447867-2
Patent No. 5447867
APPLICANT: BRIDGES, IAN/SCHUCH, WOLFGANG, GRIERSON, DONALD
TITLE OF INVENTION: RECOMBINANT DNA CONTAINING PECTIN
ESTERASE GENE SEGMENTS
NUMBER OF SEQUENCES: 4
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/24, 866
FILING DATE: 26-FEB-1993
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 720, 629
FILING DATE: 25-JUN-1991
APPLICATION NUMBER: 419, 779
FILING DATE: 29-SEP-1989
APPLICATION NUMBER: 119, 614
FILING DATE: 12-NOV-1987
SEQ ID NO: 2
LENGTH: 1636
5447867-2

Alignment Scores:

Pred. No.: 3.35e-273 Length: 1636
Score: 2390.00 Matches: 457
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatch: 0
Query Match: 100.00% Indels: 0
Gaps: 0

US-10-691-374-2 (1-457) x 5447867-2 (1-1636)

QY	1	MetValIIGLnArgpAsnSerIleuLeuLeuIleIleIlePheAAsaSerIleSer	20
DB	51	ATGGTTATTCACAAAGGATGATTCCTTCATCATTAATATTTGGCTTCATCAATTC	110
QY	21	ThrcYArgSerAsnValIleAspAsnLeuPheYcGlnValTyrAspAsnIleLeu	40

Db 111 ACTGTGAGAGCAAGTATTGATGACATTTTCAACAGCTTTATGATTAATTCCTT 170
Qy 41 GluGlnGluPheAlaHisAspPheGlnAlaTyrLeuSerTyrLeuSerLysAsn1IleGlu 60
Db 171 GAACAGAAATTTGCTCATATTTTCAAGCTTATCTTCTTATTTGAGCAAAAATATTGAA 230
Qy 61 SerAsnAsnAsn1AspLysValAspLysAsnGly1IleVal1IleAsnValLeuSer 80
Db 231 ACCAACAAATTAATTTGACAGAGTTGATTAATAATGGAAATTAAGATTAAGTACTTAC 290
Qy 81 PheGlyAlaLysGlyAspGlyLysThrTyrAspAsn1IleAlaPheGlnAlaTyrAsn 100
Db 291 TTTGAGGCTTAAGGAGTGAAGAAAAACATATGATATATTTGCAATTTGAGCAACGATG 350
Qy 101 GluAlaCysSerSerArgThrProValGlnPheValValProLysAsnLysAsnTyrLeu 120
Db 351 GAGCATGTTCTATCTAGAACACCTGTTCAATTTGTTGTTCTTAAACACAGAAATTAATCTT 410
Qy 121 LeuLysGln1IleThrPheSerGlyProCysArgSerSerLysSerValLysIlePheGly 140
Db 411 CTCAAGCAAAATCACCTTTTCAAGGTCATGACAGATCTTCTATTTGATTAAGATTTTGA 470
Qy 141 SerLeuGlnAlaSerSerLysIleSerAspTyrLysAspArgArgLeuTyrP1IleAlaPhe 160
Db 471 TCTTTAGAGCATCTAGTAAATTTCAAGACTACAAAGATGAGGCTTGGATGCTTTT 530
Qy 161 AspSerValGlnAsnLeuValAlaGlyGlyGlyThrIleAsnGlyAsnGlyGlnVal 180
Db 531 GATAGTGTTCAAAATTTAGTTGTGGAGAGAGAGAACTATCAATGGCAATGAGCAAGTA 590
Qy 181 ThrTyrProSerSerCysLys1IleAsnLysSerLeuProCysArgAspAlaProThrAla 200
Db 591 TGGTGGCCAGATCTTCTCAAAATTAATAATCACTGCGCATGAGGAGTGGACCAAGGCG 650
Qy 201 LeuThrPheTyrPheAsnCysLysAsnLeuValAsnAsnLeuLysSerLysAsnAlaGln 220
Db 651 TTAACTTCTGAGATTCGAAAATTTGAAAGTAATAATTAAGAGTAAGTAAGTAAGTAAGTA 710
Qy 221 GlnIleHisIleLysPheGlySerCysThrAsnValValAlaSerAsnLeuMetIleAsn 240
Db 711 CAATTCATATCAAAATTTGAGTCATGACATTAATGTTGAGCTTCAATTTGATGATCAAT 770
Qy 241 AlaSerAlaLysSerProAsnThrAspGlyValHisValSerAsnThrGlnTyrIleGln 260
Db 771 GCTTCAGCAAAAGCCCAATTAATCTGATGAGTCCAGTATCAATCAATCAATCAATCAAT 830
Qy 261 IleSerAspThrIleIleGlyThrGlyAspAspCysIleSerIleValSerGlySerGln 280
Db 831 ATATCTGATCTATTAATTTGGAACAGGTGATGATGATGATGATGATGATGATGATGAT 890
Qy 281 AsnValGlnAlaThrAsn1IleThrCysGlyProGlyHisGly1IleSerIleGlySerLeu 300
Db 891 AATGTGACAGCCCAAAATTAATCTGTTGTCAGAGTCAATGATTAATTAATTAATTAATTA 950
Qy 301 GlySerGlyAsnSerGlnAlaTyrValSerAsnValThrValAsnGlnAlaLysIleIle 320
Db 951 GGATCTGGAATTCAGAGCTTATGTCATTAATGATGATGATGATGATGATGATGATGAT 1010
Qy 321 GlyAlaGluAsnGlyValArgIleLysThrTyrP1GlnGlyLysSerGlyGlnAlaSerAsn 340
Db 1011 GGATGCCGAATAGAGTATGATCAAGACTTGGAGGGAGATCTGGACCAAGCTTAAGCAC 1070
Qy 341 IleLysPheLeuAsnValGluMetGlnAspValLysTyrPro1IleIleAspGlnAsn 360
Db 1071 ATCAAAATTTCTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1130
Qy 361 TyrCysAspArgValGluProCysIleGlnGlnPheSerAlaValGlnValLysAsnVal 380
Db 1131 TATGTGATCGAGTGAAGCAATGATTAACAAGTTTTCAGAGCTTCAAGTGAAGTAAGTGA 1190
Qy 381 ValTyrGluAsn1IleLysGlyThrSerAlaThrLysValAlaIleLysPheAspCysSer 400

Db 1191 GTGTATGAGAAATATCAAGGSCACAAGTGCACAAAGGTGGCCATAAATTTGATTGCAGC 1250
Qy 401 ThrAsnPheProCysGlnGlyIleIleMetGluAsn1IleAsnLeuValGlyLysGly 420
Db 1251 ACNAATCTTCCATGTGAAGAAATTAATTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 1310
Qy 421 LysProSerGlnAlaThrCysLysAsnValHisPheAsnAsnAlaGlnHisValThrPro 440
Db 1311 AAACCATCAGAGGCTACGTCGCAAAATGTCATTTTAAACAATGCTGAACATGTTACCA 1370
Qy 441 HisCysThrSerLeuGln1IleSerGluAspGlnAlaLeuLysTyrAsnTyr 457
Db 1371 CACTGCACTTCACTGAATTTTCAGAGATGAAGCTCTTTGTATTAATTAAT 1421

RESULT 2

5447867-2
/ Patent No. 5447867
/ APPLICANT: BRIDGES, IAN, SCHUCH, WOLFGANG, GRIERSON, DONALD
/ TITLE OF INVENTION: RECOMBINANT DNA CONTAINING PECTIN
/ ESTERASE GENE SEGMENTS
/ NUMBER OF SEQUENCES: 4
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/24,866
/ FILING DATE: 26-FEB-1993
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 720,629
/ FILING DATE: 25-JUN-1991
/ APPLICATION NUMBER: 419,779
/ FILING DATE: 29-SEP-1989
/ APPLICATION NUMBER: 119,614
/ FILING DATE: 12-NOV-1987
/ SEQ ID NO.2:
/ LENGTH: 1636
5447867-2

Alignment Scores:

Pred. No.: 3.35e-273 Length: 1636
Score: 2390.00 Matches: 457
Best Local Similarity: 100.00% Conservative: 0
Query Match: 100.00% Mismatches: 0
DB: Gaps: 0

US-10-691-374-2 (1-457) x 5447867-2 (1-1636)

Qy 1 MetVal1IleGlnArgAsnSer1IleLeuLeu1IleIlePheAlaSerSer1IleSer 20
Db 51 ATGATATCCAAAGGAATGATATTTCTCTTCATTAATTAATTTCTTCATCAATTTCA 110
Qy 21 ThrCysArgSerAsnVal1IleAspAsnLeuPheLysGlnValTyrAspAsn1IleLeu 40
Db 111 ACTGTGAGAGCAAGTATTGATGACATTTTCAACAGCTTTATGATTAATTCCTTCAATTTCA 170
Qy 41 GluGlnGluPheAlaHisAspPheGlnAlaTyrLeuSerTyrLeuSerLysAsn1IleGlu 60
Db 171 GAACAGAAATTTGCTCATATTTTCAAGCTTATCTTCTTATTTGAGCAAAAATATTGAA 230
Qy 61 SerAsnAsnAsn1AspLysValAspLysAsnGly1IleVal1IleAsnValLeuSer 80
Db 231 ACCAACAAATTAATTTGACAGAGTTGATTAATAATGGAAATTAAGATTAAGTACTTAC 290
Qy 81 PheGlyAlaLysGlyAspGlyLysThrTyrAspAsn1IleAlaPheGlnAlaTyrAsn 100
Db 291 TTTGAGGCTTAAGGAGTGAAGAAAAACATATGATATATTTGCAATTTGAGCAACGATG 350
Qy 101 GluAlaCysSerSerArgThrProValGlnPheValValProLysAsnLysAsnTyrLeu 120
Db 351 GAGCATGTTCTATCTAGAACACCTGTTCAATTTGTTGTTCTTAAACACAGAAATTAATCTT 410
Qy 121 LeuLysGln1IleThrPheSerGlyProCysArgSerSerLysSerValLysIlePheGly 140
Db 411 CTCAAGCAAAATCACCTTTTCAAGGTCATGACAGATCTTCTATTTGATTAAGATTTTGA 470

```
QY 141 SerLeuGluAlaSerSerLysIleSerAspTyrLysAspArgLeuTrpIleAlaPhe 160
DB 471 TCTTGAAGCATCTGTAATAATTGACATACAAAGGAGGCTTTGATGCTTTT 530
QY 161 AspSerValGlnAsnLeuValValGlyGlyGlyThrIleAsnGlyAsnGlyGlnVal 180
DB 531 GATAGCTTCAAAATTTAGTGTGTGAGAGAGAGAACTATCATATGCAATGACAAAGTA 590
QY 181 TrpTrpProSerSerCysValylIleAsnLysSerLeuProCysArgAspAlaProThrAla 200
DB 591 TGGTGCCCAAGTCTTGCAAAATAAATCAATCTCCATGACGAGATCACCAAGCGCC 650
QY 201 LeuThrPheThrAsnProCysLysAsnLeuValAsnAsnLeuLysSerLysAsnAlaGln 220
DB 651 TTAACCTTGTGAATTCGAAATTTGAAAGTGAATTAAGTAAGTAAGTAAGTAAGTAAG 710
QY 221 GlnIleHisIleLysPheGluSerCysThrAsnValValAlaSerAsnLeuMetIleAsn 240
DB 711 CAATTCATATCAAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 770
QY 241 AlaSerAlaLysSerProAsnThrAspGlyValHisValSerAsnThrGlnTyrIleGln 260
DB 771 GCTTCGCAAAAGGCCCAATATCTGATGATGATGATGATGATGATGATGATGATGAT 830
QY 261 IleSerAspThrIleIleGlyThrGlyAspAspCysIleSerIleValSerGlySerGln 280
DB 831 AATTCGATCACTATTAATGACAGGATGATGATGATGATGATGATGATGATGATGAT 890
QY 281 AsnValGlnAlaThrAsnIleThrCysGlyProGlyHisGlyIleSerIleGlySerLeu 300
DB 891 AATGTCGAGGCCCAAAATTAATCTTGTGTCGAGTCGAGTCGAGTCGAGTCGAGTCGAG 950
QY 301 GlySerGlyAsnSerGluAlaTyrValLysAsnValThrValAsnGluAlaLysIleIle 320
DB 951 GGAATCGGAAATTCAGAACCTTATGCTTAAGTAACTGTAATCAAGTAACTGTAATCA 1010
QY 321 GlyAlaGluAsnGlyValArgIleLysThrTrpGlnGlyGlySerGlyGlnAlaSerAsn 340
DB 1011 GGTGCGGAAATGGAATTCAGAACCTTGTGTCGAGTCGAGTCGAGTCGAGTCGAGTCGAG 1070
QY 341 IleLysPheLeuAsnValGluMetGluAspValLysTyrProIleIleIleAspGlnAsn 360
DB 1071 ATCAATTTCTGATGATGGAATGCAAGACGTTAAGTATCCCAATTAATTAAGCCAAAC 1130
QY 361 TyrCysAspArgValGluProCysIleGlnGlnPheSerAlaValGlnValLysAsnVal 380
DB 1131 TATGTGATCGAGTTCAGACCATGATATCAACAGTTTTCAGACGTTTCAGACGTTTCAG 1190
QY 381 ValTyrGluAsnIleLysGlyThrSerAlaThrLysValAlaIleLysPheAspCysSer 400
DB 1191 GTGTATGAAATATCAAGGCGACAGTCGACAAAGGCGCCATTAATTTGATTCAGC 1250
QY 401 ThrAsnPheProCysGluGlyIleIleMetGluAsnIleAsnLeuValGlyGluSerGly 420
DB 1251 ACAACCTTCCATGTAAGGAATTAATTAATGAGAAATATAATTTACTATGAGGAATGGA 1310
QY 421 LysProSerGluAlaThrCysLysAsnValHisPheAsnAsnAlaGlnHisValThrPro 440
DB 1311 AAACCATCAGAGGCTACGTCGCAAAATGTCATTTTACAAATGCTTACAAATGCTTAC 1370
QY 441 HisCysThrSerLeuGluLysSerGluAspGluAlaLeuLeuTyrAsnTyr 457
DB 1371 CACTGCACTTCACTAGAAATTTCAAGAGATGAAGCTCTTTGTATTAATTAAT 1421
```

```
APPLICANT: BORKHARDT, Bernard
APPLICANT: SANDER, L1111
APPLICANT: PETERSEN, Morten
APPLICANT: BUNDGARD, POULSEN, Gert
APPLICANT: BOTTERMAN, Johan
TITLE OF INVENTION: Seed Shattering
FILE REFERENCE: 2121-0138P
CURRENT APPLICATION NUMBER: US/09/051,239A
CURRENT FILING DATE: 1998-09-28
PRIOR APPLICATION NUMBER: PCT/EP96/04313
PRIOR FILING DATE: 1996-10-04
PRIOR APPLICATION NUMBER: EP 95 402241.4
PRIOR FILING DATE: 1995-10-06
PRIOR APPLICATION NUMBER: EP 95 203328.0
PRIOR FILING DATE: 1995-12-08
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO: 1
LENGTH: 1631
TYPE: DNA
ORGANISM: Brassica napus
FEATURE:
OTHER INFORMATION: Location 95-163 = region encoding the presumed
OTHER INFORMATION: endo-PG signal peptide.
OTHER INFORMATION: Location 884-900 = region of the endo-PG cDNA
OTHER INFORMATION: corresponding to oligonucleotide PG3
OTHER INFORMATION: Location 1059-1073 = region of the endo-PG cDNA
OTHER INFORMATION: complementary to oligonucleotide PG2
OTHER INFORMATION: Location 1229-1245 = region of the endo-PG cDNA
NAME/KEY: CDS
LOCATION: (95)..(1393)
OTHER INFORMATION: Location 821-837 = region of endo-PG cDNA
OTHER INFORMATION: corresponding to oligonucleotide PGI.
OTHER INFORMATION: Strain cv. Topaz.
NAME/KEY: unsure
LOCATION: (1439)
OTHER INFORMATION: n = a, c, g, t, any, other, unknown, or other
US-09-051-239A-1
Alignment Scores:
Pred. No.: 1,3e-121 Length: 1631
Score: 1113.00 Matches: 223
Percent Similarity: 65.32% Conservative: 69
Best Local Similarity: 49.89% Mismatches: 121
Query Match: 46.57% Indels: 34
DB: Gaps: 8
US-10-691-374-2 (1-457) x US-09-051-239A-1 (1-1631)
QY 7 SerIleLeuLeuLeuIleIleIlePheAlaSerSerIleSerThrCysArgSerAsnVal 26
DB 119 GCTGTTTCTTATGCGTCTTTTGTGATGCTCGCTTGCTGCAAGCTTTGAGTCAAGTA 178
QY 27 IleAspAsnLeuPheLysGlnValTyrAspAsnIleLeuGluGlnGluPheAlaHis 46
DB 179 ---GATGATGCA-----TATGCTCAT 196
QY 47 AspPheGlnAlaTyrLeuSerTyrLeuSerLysAsnIleGluSerAsnAsnIleAsp 66
DB 197 GAAGATGCA-----AGCTTCGATCCGATGATTAAATCAAGTCAACAACGACGAC 247
QY 67 LysVal-----AspLysAsnGlyIleLysValIleAsnVal 78
DB 248 GACGTTCTTACCTGAAAGCTCTGATAGACCCACTACCGAATCATCACTGTACTGTT 307
QY 79 LeuSerPheGlyAlaLysGlyAspGlyLysThrTyrAspAsnIleAlaPheGluGlnAla 98
DB 308 TCGAATCTGAGCCAAAGAGATGGAATAACCGATGATCTCAGGCTTTCAAGAAAGCA 367
QY 99 TrpAsnGluAlaCysSerSerArgThrProValGlnPheValValProLysAsnLysAsn 118
DB 368 TCGAAGAGGATGTTCAACAATGAGATTAATCTTCTTAATCTTAATCTTAATCTTAATCT 427
```

```

Qy 119 TyrlleuLysGlnIleThrPheSerGlyProCysArgSerIleSerValIle 138
Db 428 TATCTCTTAAGCTTATTAGTTCAGAGGCCATGCAATCTTACGTCTCCAGATC 487
Qy 139 PheGlySerLeuGlnIleSerIleSerIleSerIleSerIleSerIleSerIle 157
Db 488 CTAGGCACTTATACAGTCTTACAGAACGATCGATTCAGTAATGACAAACCACTG 547
Qy 158 ILeAlaPheArgSerValGlnIleValIleGlyGly-----GlyThrIleAsn 175
Db 548 CTATATTTGGAAACCTTATATATCATCAATCGATGGCGGTCCGGCGGATTTGAT 607
Qy 176 GlyAsnGlyGlnValIleThrProSerSerGlyIleAsnIleSerIleProCysArg 195
Db 608 GGAAACGGAATATCTGCTGAGCAAACTCATGCAAAATCGAACAATCTAACCAATGACA 667
Qy 196 AspAlaProThrAlaLeuThrPheThrPheCysIleAsnLeuValIleAsnLeuVal 215
Db 668 AAAGCCCAACGCTCTTACTCTCTACCAACCTTAAGATTTGATGTAAGATCTGAGA 727
Qy 216 SerIleAsnAlaGlnIleIleIleIleIleIleIleIleIleIleIleIleIle 235
Db 728 GTGAGAAATGCACAGACAGATTCAGATTCGATTGGAATGCAACAATGTTGGCGTTAAG 787
Qy 236 AsnLeuMetIleAsnAlaSerAlaIleSerProAsnThrAspGlyValIleValSerAsn 255
Db 788 AATGTTAAGATCATCTGCTCTGGCGATGTCCTCCAAACGAGATGATTCATTCATTCGCT 847
Qy 256 ThrGlnIleIleGlnIleSerAspThrIleIleGlyThrGlyAspProCysIleSerIle 275
Db 848 ACTAAAAACATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 907
Qy 276 ValSerGlySerGlnIleValIleValIleValIleValIleValIleValIleVal 295
Db 908 GAGATGATCGCAAAATGTTCAATCATGATTTACTTGGCGCCCGCTCAATGAGATC 967
Qy 296 SerIleGlySerLeuGlySerGlyAsnSerGlyAsnValIleValIleValIleValIle 315
Db 968 AGCATGGAAGCTTGGGAGATGACAAATTCGATTCGATTCGATTCGATTCGATTCGATTC 1027
Qy 316 GlnValIleIleIleGlyAlaGlnValIleValIleValIleValIleValIleValIle 335
Db 1028 GGTCTCGCTCTGAGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1087
Qy 336 GlyGlnAlaSerAsnIleIleIleIleIleIleIleIleIleIleIleIleIleIle 355
Db 1088 GGAAGTCTTAAGAACTTAATTCGAAACATTCGATTCGATTCGATTCGATTCGATTCGATTC 1147
Qy 356 IleIleAspGlnAsnIleIleIleIleIleIleIleIleIleIleIleIleIleIle 375
Db 1148 ATATGACCCAGAACTACGCGCAAG--GACAAATGCAACACCAAGATCTCGCGTT 1204
Qy 376 GlnValIleAsnValIleIleIleIleIleIleIleIleIleIleIleIleIleIle 395
Db 1205 CAAGTGAACATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1264
Qy 396 LysPheAspCysSerThrAsnIleIleIleIleIleIleIleIleIleIleIleIleIle 415
Db 1265 ATGTTAATGAGTGAATATATCCAAAGGATTTGCTGGAATGATGATGATGATGATGATGAT 1324
Qy 416 ValGlyGlySerGlyLysProSerGlnIleThrCysIleAsnValIleIlePheAsnAla 435
Db 1325 AAAGCA-----GGAAGACCTTCTCGCAAAATGCAATGTTAAGGATAAA 1369
Qy 436 GlnIleValIleThrProIleCys 442
Db 1370 GGCACCTTTCTCCCTAAATGC 1390

```

RESULT 4
 US-10-151-668-1
 ; Sequence 1, Application US/10151668
 ; Patent No. 6797861

```

; GENERAL INFORMATION:
; APPLICANT: ULVSKOV, Peter
; APPLICANT: CHILD, Robin
; APPLICANT: VAN ONCKELIN, Henri
; APPLICANT: PRINSEN, Els
; APPLICANT: BORKHARDT, Bernard
; APPLICANT: SANDER, Lilij
; APPLICANT: PETERSEN, Morten
; APPLICANT: BONDARD, POTSEN, Gert
; APPLICANT: BOTTERMAN, Johan
; TITLE OF INVENTION: Seed Shattering
; FILE REFERENCE: 2121-0138P
; CURRENT APPLICATION NUMBER: US/10/151,668
; PRIOR FILING DATE: 2002-05-21
; PRIOR APPLICATION NUMBER: US/09/051,239
; PRIOR FILING DATE: 1998-09-28
; PRIOR APPLICATION NUMBER: PCT/EP96/04313
; PRIOR FILING DATE: 1996-10-04
; PRIOR APPLICATION NUMBER: EP 95 402241.4
; PRIOR FILING DATE: 1995-10-06
; PRIOR APPLICATION NUMBER: EP 95 203328.0
; PRIOR FILING DATE: 1995-12-08
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1631
; TYPE: DNA
; ORGANISM: Brassica napus
; FEATURE:
; OTHER INFORMATION: Location 95-163 = region encoding the presumed
; OTHER INFORMATION: endo-PG signal peptide.
; OTHER INFORMATION: Location 884-900 = region of the endo-PG cDNA
; OTHER INFORMATION: corresponding to oligonucleotide PG3
; OTHER INFORMATION: Location 1059-1073 = region of the endo-PG cDNA
; OTHER INFORMATION: complementary to oligonucleotide PG2
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (95)..(1393)
; OTHER INFORMATION: Location 1229-1245 = region of the endo-PG cDNA
; OTHER INFORMATION: complementary to oligonucleotide PG5
; OTHER INFORMATION: Location 821-837 = region of endo-PG cDNA
; OTHER INFORMATION: corresponding to oligonucleotide PG1.
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (11439)
; OTHER INFORMATION: Strain cv. Topaz.
; OTHER INFORMATION: n = a, c, g, t, any, other, unknown, or other
; US-10-151-668-1

Alignment Scores:
Pred. No.: 1,3e-121 Length: 1631
Score: 1113.00 Matches: 223
Percent Similarity: 65.32% Conservative: 69
Best Local Similarity: 49.89% Mismatches: 121
Query Match: 46.57% Indels: 34
Gaps: 8

US-10-691-374-2 (1-457) x US-10-151-668-1 (1-1631)
Qy 7 SerIleLeuLeuLeuIleIleIlePheAlaSerSerIleSerThrCysArgSerAsnVal 26
Db 119 GCTGTTTCTTAAGCTTATTAGTTCAGAGGCCATGCAATCTTACGTCTCCAGATC 178
Qy 27 IleAspPheAsnLeuPheGlnValIleIlePheAsnIleLeuGlnIlePheAlaIle 46
Db 179 ---GATGATGCA-----TATGTCAT 196
Qy 47 AspPheGlnAlaIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 66
Db 197 GAAGATGGA-----AGCTTCGATCCGATGATGATGATGATGATGATGATGATGATGAT 247
Qy 67 LysVal-----AspLysAsnGlyIleLeuValIleAsnVal 78

```


Db 248 GAAGTTCTTACCTTGAAAGCTCTGATAGACCCACCTACCGAATCATCACTGTACTGTT 307
Qy 79 LeuSerPheGlyAlaIysGlyAspGlyIysThrTyAspAsnIleAlaPheGluGlnAla 98
Db 308 TCGAAGCTTCGAGCCCAAGAGATGAGAAAAACCGAATGATCACTGAGCTTCAAGAAAGCA 367
Qy 99 TrpAsnGlnAlaCysSerSerArgThrProValGlnPheValValProIysAsnIysAsn 118
Db 368 TGGAGAGAGACGATGTTCAACAAATGAGATTACTTCTTAATTCCTTAAGGAAAGACT 427
Qy 119 TyLeuLeuLysGlnIleThrPheSerGlyProCysArgSerSerIleSerValIysIle 138
Db 428 TATCTCTTAAGTATTAGATTCAAGGCCCATGGAATCTTAGTGCTTCGACATC 487
Qy 139 PheGlySerLeuGlnAlaSerSerIysIleSerAspTyr--LysAspArgArgLeuTyr 157
Db 488 CTAGGCACTTATACACTTCTACAAACCATCGAATTACAGTAAATGACAGAACCACTGC 547
Qy 158 IleAlaPheAspSerValGlnAsnLeuValValGlyGly-----GlyThrIleAsn 175
Db 548 CTTATTTTGGAGACGCTTAATAATCTATCAATCGATGCGCGCGGATGTGTGAT 607
Qy 176 GlyAsnGlyGlnValTrpTrpProSerSerCysIysIleAsnIysSerLeuProCysArg 195
Db 608 GCGAAGCGAAATATCTGCTGCGCAAACTATGCAAAATGACAAATCTTAAGCCATCACA 667
Qy 196 AspAlaProThrAlaLeuThrPheTrpAsnCysIysAsnLeuIysValAsnLeuIys 215
Db 668 AAGCGCCCAACGGCTCTTACTCTCTACAACTTAAGAAATTTGAAATGGAATCTGAGA 727
Qy 216 SerIysAsnAlaGlnGlnIleHisIleIysPheGlyIysCysThrAsnValAlaIleSer 235
Db 728 GTGAGAAATGACAGAGATTCAGATTCGATGGAATAATGCAACATCTTGCGCTTAAG 787
Qy 236 AsnIleMetIleAsnAlaSerAlaIysSerProAsnThrAspGlyValHisValSerAsn 255
Db 788 AATGTTAAGATCACTGCTCTGCGCATAGTCCCAACAGGATGCTATCATTCGTTGCT 847
Qy 256 ThrGlnTyrIleGlnIleSerAspThrIleIleGlyThrGlyAspAspCysIleSerIle 275
Db 848 ACTAAATAATTCGAATCTCCAAATTCAGACATGGAGAGTGATATGTATATCATAT 907
Qy 276 ValSerGlySerGlnAsnValGlnAlaThrAsnIleThrCysGlyProGlyHisGlyIle 295
Db 908 GAGGATGATCGAAAAATGTTCAATCAATGATTAATCTGCGGCCGCTCATGGGATC 967
Qy 296 SerIleGlySerLeuGlySerGlyAsnSerGlnAlaTyrValSerAsnValThrValAsn 315
Db 968 AGCATTTGGAAGCTTGGGGATGACAAATTCGAAGCTTATGTATCGGGAATGATGTGAT 1027
Qy 316 GluAlaLysIleIleGlyAlaGluAsnGlyValArgIleIysThrTrpGlnGlyIysSer 335
Db 1028 GGTGCTACCTCTCTGAGACTGACATGAGTGAATCAAGACTTACCAAGGAGGGGTCA 1087
Qy 336 GlyGlnAlaSerAsnIleIysPheLeuAsnValGluMetGlnAspValLysTyrProIle 355
Db 1088 GGAAGCTCTGAAGAACCTTAATTCAAAACATTCGATGATATATCAAGAAATCCGATC 1147
Qy 356 IleIleAspGlnAsnTyrCysAspArgValGluProCysIleGlnGlnPheSerAlaVal 375
Db 1148 ATATATGACCAAGAACTACGCGCAAG---GACAAATGCGAACAAGAAATCTGCGGCT 1204
Qy 376 GlnValLysAsnValValIysGluAsnIleIysGlyThrSerAlaThrLysValAlaIle 395
Db 1205 CAAGTGAACAATGTCGTATATCAGAACATAAAGATACAGACGCAACAATGTGGCGAT 1264
Qy 396 LysPheAspCysSerThrAsnPheProCysGluGlyIleIleMetGluAsnIleAsnLeu 415
Db 1265 ATGTTATTTGCGATGTGAATTCATGCAAGGATATGCTTATGAGATGTGAACATC 1324
Qy 416 ValGlyIysSerGlyLysProSerGlnAlaThrCysIysAsnValHisPheAsnAsnAla 435
Db 1325 AAGAG-----GGAAAGCTTCTTGGGAAATGTCAATGTTAAGATTA 1369

Qy 436 GluHisValThrProHisCys 442
Db 1370 GGCAGCTTTCTCTTAATATGC 1390

RESULT 5
US-08-941-532-5
Sequence 5, Application US/08941532
Patent No. 6096946
GENERAL INFORMATION:
APPLICANT: ROBERTS, Jeremy Alan
APPLICANT: COUP, Simon Allan
APPLICANT: JENKINS, Elizabeth Sarah
TITLE OF INVENTION: CONTROL OF POD DENSIFICATION
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/941,532
FILING DATE: 30-SEP-1997
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/GB96/00757
FILING DATE: 29-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9506684.1
FILING DATE: 31-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Biamond, Robert W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0623.0580001/RWE
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1657 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 145..1446
US-08-941-532-5

Alignment Scores:
Pred. No.: 2.05e-120
Score: 1103.00
Percent Similarity: 65.54%
Best Local Similarity: 49.77%
Query Match: 46.15%
DB: 3
Gaps: 7

US-10-691-374-2 (1-457) x US-08-941-532-5 (1-1657)

Qy 7 SerIleLeuLeuLeuIleIleIlePheAlaSerSerIleSerThrCysArgSerAsnVal 26
Db 169 GCTATTTCTTATGCGTTCTTGTGATGCTCGCTGCGCAAGCTTTGATGATGCAACGTA 228
Qy 27 IleAsp-----AspAsnLeuPheLysGlnValTyrAspAsnIleLeuGln 41
Db 229 GATGATGATATGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 282

```

QY 42 GlnGluPheAlaHisAspPheGlnAlaTyrLeuSerTyrLeuSerIleAsnIleGluSer 61
Db 283 CTCAACACGACGACGACGCTTCTTACCTGAAAACTCCGATGACCCACTACCGAATCA 342
QY 62 AsnAsnAsnIleAspIleValAspIleAsnGlyIleValIleAsnValIleuSerPhe 81
Db 343 TCA-----ACTGTTGTTGTTTCCAACTTC 366
QY 82 GlnAlaIleGlyAspGlyIleValTyrAspAsnIleAlaPheGluGlnAlaTyrAsnGlu 101
Db 367 GAGCAAAAGGTATGAGAAAAACCGATGATACAGGCTTTCAGAAACACGAGAAAG 426
QY 102 AlaCysSerSerArgThrProValGlnPheValProIleAsnIleAsnValIleuSer 121
Db 427 GCATGTTTCACAAATGAGAGTACTTCTTATCTTAAAGGAGAGCTTATCTCTT 486
QY 122 LysGlnIleThrPheSerGlyProCysArgSerSerIleSerValIlePheGlySer 141
Db 487 AACTCTATTAGATTCAAGGCGCATCAATTCATTACGTTCAATCTCAATCTTATTTG 546
QY 142 LeuGlnAlaSerSerIleSerAspTyr--LysAspArgLysLeuTyrIleAlaPhe 160
Db 547 TTATCACTTCTTACAAACGATCGATTAAGATGACAAACACGCTTATTTG 606
QY 161 AspSerValGlnAsnLeuValGlyIleGly-----GlyThrIleAsnGlyAsnGly 178
Db 607 GAGGACGTTAATATCTAATCATGATGCGGCTCGCGGCGGATTTGATGAGCAACGA 666
QY 179 GlnValIleTyrProSerSerCysIleValIleAsnIleSerLeuProCysArgAspAlaPro 198
Db 667 AAAATCTGTCGCAAAACTCATGCAAAATCGACAAATCTAAGCCATCAACAAAGCCCA 726
QY 199 ThrAlaLeuThrPheTyrAsnCysIleAsnLeuValAsnAsnLeuLysSerIleAsn 218
Db 727 ACGGCTCTTACTCTACACCTTAACATTGATGAAAGATCTGAGAGTGAAGAAAT 786
QY 219 AlaGlnGlnIleHisIleValPheGluSerCysThrAsnValAlaIleSerAsnLeuMet 238
Db 787 GCACGACGATTCAGATTCATGATGAGAAATGCAACAGTGTATGATTAAGATGTAAAG 846
QY 239 IleAsnAlaSerAlaLysSerProAsnThrAspGlyValHisValSerAsnThrGlnTyr 258
Db 847 ATCACTGCTCTCGGCGATGATCCCAACGCGATCGATTTATATCTTCTCTTAAAC 906
QY 259 IleGlnIleSerAspThrIleIleGlyThrGlyAspAspCysIleSerIleValSerGly 278
Db 907 ATTGAAATCTCAATTCAGACATTTGGACAGGTGATGATGATTCATTCAGGATGA 966
QY 279 SerGlnAsnValGlnAlaThrAsnIleThrCysGlyProGlyHisGlyIleSerIleGly 298
Db 967 TCCCAAAATGTTCAAAATCATATGATTTAATTCGGCCCGCGCATGCGCATGCGATTTGA 1026
QY 299 SerLeuGlySerCysIleAsnSerGlnAlaTyrValSerAsnValThrValAsnGlnAlaLys 318
Db 1027 AGCTTGGGGGATCACAATTCAAAGCTTATGATCGGAAATTAATGATGATGCTTACG 1086
QY 319 IleIleGlyAlaGlnAsnGlyValArgIleLysThrTyrPheGlnGlySerGlyGlnAla 338
Db 1087 CTCTGAGACGTGACATGAGATTAAGATCAACATCTTACAGGAGGCGTCAAGAACTGCT 1146
QY 339 SerAsnIleLysPheLeuAsnValGlnMetGlnAspValLysTyrProIleIleIleAsp 358
Db 1147 AAGAAATTAATTAATTCAAACATTCGTATGATATGTCAGAAATCCGATCATTAATGAC 1206
QY 359 GlnAsnTyrCysAspArgValGlnProCysIleGlnGlnPheSerAlaValGlnValLys 378
Db 1207 CAGAACTACTGGGACAG--GACAAATCCAAACAAAGAAATCTGGGCTTCAAGTGAAAC 1263
QY 379 AsnValIleTyrGlnAsnIleLysGlyThrSerAlaThrLysValAlaIleLysPheAsp 398
Db 1264 AATGTCGTATCGGAACATCAAGGTACGAGCCCAAGCATGTGGGATTAATGTTTAAT 1323

```

```

QY 399 CysSerThrAsnProCysGlnGlyIleIleMetGluAsnIleAsnLeuValGlyGlu 418
Db 1324 TGCAGTGTGAATATCATCTCCAAAGTATTTGCTTGAAATGTGAATCAAAAGCA-- 1380
QY 419 SerCysIleProSerGlnAlaThrCysIleAsnValHisPheAsnAsnAlaGlnHisVal 438
Db 1381 -----GAAAGCTTCTTCAAAATGTCAAATGTTAAGATTAAGCAACCGTT 1428
QY 439 ThrProHisCys 442
Db 1429 TCTCTAATATGC 1440

RESULT 6
US-08-467-023-133
; Sequence 133, Application US/08467023
; Patent No. 6090386
; GENERAL INFORMATION:
; APPLICANT: Griffith, Irwin J.
; APPLICANT: Pollock, Joanne;
; APPLICANT: Bond, Julian F.;
; APPLICANT: Garmann, Richard D.;
; APPLICANT: Kuo, Wei-Chang;
; APPLICANT: Jeung, Su-mei H.;
; APPLICANT: Brauer, Andrew;
; APPLICANT: Exley, Mark A.;
; APPLICANT: Powers, Steven P.
; TITLE OF INVENTION: Allergenic Proteins And Peptides From
; TITLE OF INVENTION: Japanese Cedar Pollen
; NUMBER OF SEQUENCES: 261
; CORRESPONDENCE ADDRESSES:
; ADDRESSER: Immunologic Pharmaceutical Corporation, Inc.
; STREET: 610 Lincoln St
; CITY: Waltham
; STATE: MA
; COUNTRY: USA
; ZIP: 02154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,023
; FILING DATE: June 6, 1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/350,225
; FILING DATE: December 6, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jane E. Remillard
; REGISTRATION NUMBER: 38,872
; REFERENCE/DOCKET NUMBER: 025.6 USD2 (IMI-028CPD2)
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 133:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1726 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 42..1586
; US-08-467-023-133

Alignment Scores:
Pred. No.: 7.15e-83 Length: 1726
Score: 787.00 Matches: 168
Percent Similarity: 58.07% Conservative: 73
Best Local Similarity: 40.48% Mismatches: 160
Query Match: 32.93% Indels: 14

```

```

DB: 3 Gaps: 7
US-10-691-374-2 (1-457) x US-08-467-023-133 (1-1726)
QY 45 AAlNHASpPhgElnAlaTyrLeu-----SerTyrLeuSerIysAsnIleGluSer 61
DB 105 GCGAAGATCAATCTGCCAAATATTATGTTGACAGCTTGTGCAAAAATATCTTAATG 164
QY 62 AAsnAsnIleAspIysValAsp-----LysAsnGlyIleIysValIleAsnVal 78
DB 165 AATCGAGATTAAAGAAAGTTGAGCATCTCTCATGATGCTTAACAACCTTCAATG 224
QY 79 LeuSerPhgElnAlaIysGlyAspGlyIysThrTyrAspAsnIleAlaPhgElnAla 98
DB 225 GAAAGATATGGCCGATGGAGGAAAGCATGATTCACCTGAGCAATTTTCAACAGCA 284
QY 99 TTPAsnGlnAlaCysSerSerArgThrProValGlnPhgValAlaProIysAsnIysAsn 118
DB 285 TGGCAAGCTGCATGCAAAAACCATCA---GCAATGTGCTTGTGCGACGACAGCAAGAA 341
QY 119 TyrLeuLeuIysGlnIleThrPheserGlyProCysArgSerIleSerValIysIle 138
DB 342 TTTGTGTAAACAATCTGTCTTCAATGGCCATGTCACCTTCTTTAAGTA 401
QY 139 PhgElnSerLeuGlnIleAspSerIysIleSerAspTyrIysAspArgIleuTyrIle 158
DB 402 GATGGATATATACCTGCTGACCAAAATCCAGCGAGCTGGAGAAATATAGATATGGTGG 461
QY 159 AAlaPhaSpSerValGlnAsnLeuValIleGlyIleGlyIleThrIleAsnGlyAsnGly 178
DB 462 CAGTTTGCTAAACTTAACAGTTTATCTTAATGGGTAAAGGTAAATGATGGCAAGGA 521
QY 179 GlnValIleTyrProSerSerCysIys---IleAsnIysSerLeuProCys-----Arg 195
DB 522 AAACAATGGGGCTGGCCCATGTAAATGGGTCAATGAGACAGAAATTTGCAACGATCG 581
QY 196 AAspAlaProThrAlaLeuThrPheTyrAsnCysValAsnLeuIysValAsnLeuIys 215
DB 582 GATAGACCAACACCTTAATTCGATTTTCCACGGGCTGTATATTCAGACGATGAA 641
QY 216 SerIysAsnAlaGlnIleIleIleIleIleIleIleIleIleIleIleIleIleIle 235
DB 642 CTAAATGACAGTCCGCAATTTCTAGTTTGGGAATTTGAGGAGGAAATCAATC 701
QY 236 AsnLeuMetIleAsnAlaSerAlaIysSerProAsnThrAspGlyValIleValSerAsn 255
DB 702 GGCATTAGATATACGGCAGCAGACAGACAGCTTAACATGATGATATCTTGGCA 761
QY 256 ThrGlnTyrIleGlnIleSerAspThrIleIleGlyThrGlyAspAspCysIleSerIle 275
DB 762 TCTAAAACTTTCACTTAACAAAAGAACAGATGGAACAGGGGATACGCGCTATA 821
QY 276 ValSerGlySerGlnAsnValGlnAlaThrAsnIleThrCysGlyProGlyIleGlyIle 295
DB 822 GGGACAGGGGCTCTTAATATTTGATGAGATCTGATTTGGCGTCCAGCCATGGAAT 881
QY 296 SerIleGlySerLeuGlySerGlyAsnSerGlyAlaIysValSerAsnValThrValAsn 315
DB 882 AGTATAGGAAGTCTTGGGAGGAAACCTAGACAGAGTTTCATACCTGCGACCTAAAT 941
QY 316 GlnAlaIysIleIleGlyAlaGlnAsnGlyValArgIleIleIleIleIleIleIleIle 335
DB 942 GGGGCTTAATTTCACTTAACAAAAGAAATGATTAAGATCAAAACATGCGAGGTGTCA 1001
QY 336 GlnGlnAlaSerAsnIleIysPheLeuAsnValGlnMetGlnAspValIysTyrProIle 355
DB 1002 GGCATGGCAAGCATATTAATTTATGAGAAATGTTGAATTAATTCGAGAAACCCATA 1061
QY 356 IleIleAspGlnAsnTyrCysAspArgValGlnProCysIleGlnIlePheSerAlaVal 375
DB 1062 TTAATTAATCAATTTCACTGACCTTCTGCTTGGCAAAACAGAGGTCTGCGGAT 1121
QY 376 GlnValIysAsnValIleTyrGlnAsnIleIleGlyThrSerAlaThrIysValAlaIle 395

```

```

DB 1122 CAATCCAGATGTGACATACAGAAACATACCTGGACATCGACACAGCAGCAATT 1181
QY 396 LysPheAspCysSerThrAsnPhgProCysGlnGlyIleIleMetGlnAsnIleAsnLeu 415
DB 1182 CAACCTTAAGTGCAGTACAGATGACCTGCAAGATATGAAGTATAGTATATCTTGG 1241
QY 416 ValGlyIysSerGlyIysProSerGlnAlaThrCysIysAsnValIlePheAsn---- 434
DB 1242 AAGCTTACCTCAGGAA-----ATTGCTTCTGCTTAATGATTAATGCAAAATGATAT 1295
QY 435 ---AlaGlnIleValIleThrProIleCysThrSerLeuGlnIleSer 448
DB 1296 TTCAGTGCACAGCTCATCCTGCTGACAGAAATTTAATGTCAGT 1340

RESULT 7
US-08-467-023-141
Sequence 141, Application US/08467023
Patent No. 6090366
GENERAL INFORMATION:
APPLICANT: Griftech, Irwin J.;
APPLICANT: Pollock, Joanne;
APPLICANT: Bond, Julian F.;
APPLICANT: Garman, Richard D;
APPLICANT: Kuo, Mei-Chang;
APPLICANT: Yeung, Siu-mei H.;
APPLICANT: Brauer, Andrew;
APPLICANT: Exley, Mark A.;
APPLICANT: Powers, Steven P.
TITLE OF INVENTION: Allergenic Proteins And Peptides From
TITLE OF INVENTION: Japanese Cedar Pollen
NUMBER OF SEQUENCES: 261
CORRESPONDENCE ADDRESS:
ADDRESS: Immunologic Pharmaceutical Corporation, Inc.
STREET: 610 Lincoln St
CITY: Waltham
STATE: MA
COUNTRY: USA
ZIP: 02154
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,023
FILING DATE: June 6, 1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/350,225
FILING DATE: December 6, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Jane E. Remillard
REGISTRATION NUMBER: 38,872
REFERENCE/DOCKET NUMBER: 025.6 USD2 (IMI-028CPD2)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 141:
SEQUENCE CHARACTERISTICS:
LENGTH: 1479 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-467-023-141

Alignment Scores:
Pred. No.: 6,33e-83
Score: 786.50
Percent Similarity: 58.66%
Best Local Similarity: 40.84%
Query Match: 32.91%
Length: 1479
Matches: 165
Conservative: 72
Mismatch: 156
Indels: 11

```

DB: 3 Gaps: 6

US-10-691-374-2 (1-457) x US-08-467-023-141 (1-1479)

QY 53 SerTyreSerIysAsnIleGluSerAsnAsnIleAspIysValAsp----- 69

DB 31 AGTGTGCGAATAATCTTAGATCGAATCGAGTTTAAAGATGAGCTTCTGCT 90

QY 70 LysAsnGlyIleValIleAsnValLeuSerPheGlyValAlaGlyIysThr 89

DB 91 CATATGCTATCAACATCTTCAATGGAAGTAAAGCGCGAGTAGCGCATGGAAGCAT 150

QY 90 TyrAspAsnIleAlaPheGluGlnAlaTyrAsnGluAlaCysSerSerArgThrProVal 109

DB 151 GATTGCACTGAGCATTTTCAACAGCATGSCAAGCTGCATGCAAAACCATCA--GCA 207

QY 110 GlnPheValValProLysAsnLysAsnTyrLeuLeuYsgIleThrPheSerGlyPro 129

DB 208 ATGTTGCTTGTGCGACGACGAGCAAAATTTGTTGAACAAATCTTCTTCAATGGGCA 267

QY 130 CysArgSerSerIleSerValIysIlePheGlySerLeuGlnAlaSerSerIleSer 149

DB 268 TGTCAACCTCACTTCTTAAAGTAGATGAGTGGGATTAATAGCTGCTACCAAAATCCAGCG 327

QY 150 AspIlyLysAspArgArgLeuTyrIleAlaPheAspSerValGlnAsnLeuValIleGly 169

DB 328 AGCTGGAAGATAATGATATGATGTTGCAAGTTTGTCTAACTTACAGGTTTACTCTAATG 387

QY 170 GlyIleGlyIleThrIleAsnGlyAsnGlyGlnValTyrTyrProSerSerCysIys--Ile 188

DB 388 GGTAAAGGTGTATTGATGGGCAAGAAACAAATGGTGGCTGCGCAATGTAATGGCTC 447

QY 189 AsnLysSerLeuProCys-----ArgAspAlaProThrIleLeuThrPheTyrPancCys 206

DB 448 AATGACAGCAAAATTTGCAACGATCGTATGACCAACAGCATTAATTCGATTTTTC 507

QY 207 LysAsnLeuLysValAsnAsnLeuLysSerLysAsnIleGlnIleHisIleLysPhe 226

DB 508 ACGGCTGTGATTAATCCAAAGCATGAACTAATGAACTCCCAATTTCAATTAAGTTT 567

QY 227 GluSerCysThrAsnValAlaIleSerAsnLeuMetIleAsnAlaSerAlaLysSerPro 246

DB 568 GGGATGTGTGAGGAGTAATAATCATCGCATTAAGTATTAACGCAACGAGACAGCTCT 627

QY 247 AsnThrAspGlyValHisValSerAsnThrGlnTyrIleGlnIleSerAspThrIleIle 266

DB 628 AACCTGATGGAATTTGATATCTTTCATTAATAAATCTTCACTTAACAAAGAACACGATA 687

QY 267 GlyThrGlyAspAspCysIleSerIleValSerGlySerGlnAsnValGlnAlaThrAsn 286

DB 688 GGAACGAGGATGATCGCTCGCTAATAGCACAGGGCTTCTTAATATGTCATGAGAT 747

QY 287 IleThrCysGlyProGlyHisGlyIleSerIleGlySerLeuGlySerGlyAsnSerGlu 306

DB 748 CTGATTTGGCGGTCAAGCCATGGAATTAAGTAAGAGCTTGGAGGAAACCTTGA 807

QY 307 AlaTyrValSerAsnValThrValAsnGlnAlaLysIleGlyValAlaGluAsnGlyVal 326

DB 808 GCGAGGTTTCATACGTCACGTAATATGGGCTTAATTCATAGACACCAAAATGATTA 867

QY 327 ArgIleLysThrTyrGlnGlyIleSerGlyGlnAlaSerAsnIleLysPheLeuAsnVal 346

DB 868 AGAATCAAAACATGCAAGGTGTGTCAGGATGCGAAGCCATTAATTAATTAATGAGAAATGTT 927

QY 347 GluMetGlnAspValLysTyrProIleIleIleAspGlnAsnTyrCysAspArgValGlu 366

DB 928 GAAATGAATTAATTCGAGAACCCCATATTAATTAATCACTTCACTCACTTCACTTCT 987

QY 367 ProCysIleGlnGlnPheSerAlaValGlnValLysAsnValValTyrGluAsnIleLys 386

DB 988 GCTTGCCAAAACAGAGGTGTGCGCTCAATCCAAAGATGTGACATTAAGAACATACGT 1047

QY 387 GlyThrSerAlaThrLysValAlaIleLysPheAspCysSerThrAsnThrProCysGlu 406

DB 1048 GGGACATCAGCAACAGCAGCAGCATTTCAACTTAAGTGCAGTATGCTTGCAGAA 1107

QY 407 GlyIleIleLeuGlnIleAsnLeuValAlaGlyLysGlyLysProSerGluAlaThr 426

DB 1108 GATTAATAAGCTAAGTATTAATCTTGAAGCTTACTTCAAGGAGAA-----ATTGCTTCC 1161

QY 427 CysLysAsnValHisPheAsnAsn-----AlaGluHisValThrProHisCysThrSer 444

DB 1162 TGCTTAATGATTAATGCAATGAGATATTTCACTGAGCAACGTCATCCCTGCATGAGAT 1221

QY 445 LeuGluIleSer 448

DB 1222 TTAAGTCCAACT 1233

RESULT 8

US-08-467-023-140

/ Sequence 140, Application US/08467023

/ Patent No. 6090386

/ GENERAL INFORMATION:

/ APPLICANT: Griffith, Irwin J.;

/ APPLICANT: Pollock, Joanne;

/ APPLICANT: Bond, Julian F.;

/ APPLICANT: Garman, Richard D.;

/ APPLICANT: Kuo, Mei-Chang;

/ APPLICANT: Yeung, Siu-mei H.;

/ APPLICANT: Brauer, Andrew;

/ APPLICANT: Ekley, Mark A.;

/ APPLICANT: Powers, Steven P.

/ TITLE OF INVENTION: Allergenic Proteins And Peptides From

/ NUMBER OF SEQUENCES: 261

/ CORRESPONDENCE ADDRESS:

/ ADDRESS: Immunologic Pharmaceutical Corporation, Inc.

/ STREET: 610 Lincoln St

/ CITY: Walham

/ STATE: MA

/ COUNTRY: USA

/ ZIP: 02154

/ COMPUTER READABLE FORM:

/ MEDIUM TYPE: Floppy disk

/ COMPUTER: IBM PC compatible

/ OPERATING SYSTEM: PC-DOS/MS-DOS

/ SOFTWARE: Patentin Release #1.0, Version #1.25

/ CURRENT APPLICATION DATA:

/ APPLICATION NUMBER: US/08/467.023

/ FILING DATE: June 6, 1995

/ CLASSIFICATION: 424

/ PRIORITY APPLICATION DATA:

/ APPLICATION NUMBER: 08/350.225

/ FILING DATE: December 6, 1994

/ ATTORNEY/AGENT INFORMATION:

/ NAME: Jane E. Remillard

/ REGISTRATION NUMBER: 38,872

/ TELECOMMUNICATION INFORMATION:

/ TELEPHONE: (617) 227-7400

/ TELEFAX: (617) 227-5941

/ INFORMATION FOR SEQ ID NO: 140:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 1395 base pairs

/ TYPE: nucleic acid

/ STRANDEDNESS: single

/ TOPOLOGY: linear

/ MOLECULE TYPE: CDNA

/ US-08-467-023-140

Alignment Scores:

Pred. No.: 9.02e-81

Score: 768.00

Percent Similarity: 59.01%

Best Local Similarity: 41.51%

Query Match: 32.13%

Length: 1395

Matches: 159

Conservative: 67

Mismatches: 149

Indels: 8

[illegible]

```

Db      1027 ATAAAGCTAAGTATATATCTTTCAAGCTTACCTGACAGGAAA-----ATTGCTTCTGC 1080
Cy      428 LyxanValHispheAnaen-----AlaGluHisValIhrProHisCyGtrhSerLeu 445
Db      1081 CTTAAATATATATGCAAAATGATATTTTCAGTGACACGTCATCCCTGCATCAAGAAATTTA 1140
Cy      446 GluLeSer 448
Db      1141 AGTCCAAGT 1149

RESULT 9
US-08-467-023-139
Sequence 139, Application US/08467023
Patent No. 6090386
GENERAL INFORMATION:
APPLICANT: Griffith, Irwin J.;
APPLICANT: Pollock, Joanne;
APPLICANT: Bond, Julian F.;
APPLICANT: Garman, Richard D;
APPLICANT: Kuo, Mei-Chang;
APPLICANT: Yeung, Siu-mei H.;
APPLICANT: Brauer, Andrew;
APPLICANT: Exley, Mark A.;
APPLICANT: Powers, Steven P.
TITLE OF INVENTION: Allergenic Proteins and Peptides From
NUMBER OF INVENTION: Japanese Cedar Pollen
CORRESPONDENCE ADDRESS: 261
ADDRESS: Immunologic Pharmaceutical Corporation, Inc.
STREET: 610 Lincoln St
CITY: Waltham
STATE: MA
COUNTRY: USA
ZIP: 02154
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,023
FILING DATE: June 6, 1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/350,225
FILING DATE: December 6, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Jane E. Remillard
REGISTRATION NUMBER: 38,872
REFERENCE/DOCKET NUMBER: 025.6 USD2 (IMI-028CPD2)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 139:
SEQUENCE CHARACTERISTICS:
LENGTH: 1410 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-467-023-139

Alignment Scores:
Pred. No.: 9,18e-81 Length: 1410
Score: 768.00 Matches: 159
Percent Similarity: 59.01% Conservative: 67
Best Local Similarity: 41.51% Mismatches: 149
Query Match: 32.13% Indels: 8
DB: 3 Gaps: 5

US-10-691-374-2 (1-457) x US-08-467-023-139 (1-1410)

```

```

Qy 71 AsnGlyIleuValIleAsnValLeuSerPheGlyValAluSerGlyAspGlyIleThrTy 90
Db 25 GATCTATCAACATCTTCAATGTGAAAGATATGGCGCATGAGCAATGAAAGCATGAT 84
Qy 91 AspAsnIleAlaPheGluGlnAlaIlePheGlnAlaCysSerSerArgThrProValGln 110
Db 85 TGCACTGAGGCATTTTCAACAGCATGCGCAAGCTGATGCAAAACCATCA---GCATG 141
Qy 111 PheValValProValAsnValAsnValLeuValGlnIleThrPheSerGlyProCys 130
Db 142 TTGCTTGTCCAGCAGCAAGAAATTTGTTAAACATCTGTTCAATGCGGCATGT 201
Qy 131 ArgSerSerIleSerValIlePheGlySerLeuGlnAlaSerSerIleSerArg 150
Db 202 CAACCTCACTTCTTTAAGTAGATGGATTAATAGCTGCTTACCAAAATCCAGCGGC 261
Qy 151 TyrLeuAspArgArgLeuThrIleAlaPheAspSerValGlnAsnLeuValAluGly 170
Db 262 TGGAAAGATTAATTAATGATATGTTGCAATTTGCTAAACTTACAGTTTACTCTATG 321
Qy 171 GlyGlyThrIleAsnGlyAsnGlyValIleThrProSerSerCysGly---IleAsn 189
Db 322 AAAGGTGATATGATGGGCAAGAAACATATGCTGGCTGGCAATGTAATGGGTCAAT 381
Qy 190 LysSerLeuProCys-----ArgAspAlaProThrAlaLeuThrPheTrpAsnCys 207
Db 382 GCACAGAAATTTGCAACGATCGATGATAGACCAACAGCATTAATTCATTTTCCAG 441
Qy 208 AsnLeuLysValAsnAsnLeuLysSerLysAsnAlaGlnGlnIleHisIleLysPheGlu 227
Db 442 GGTCTGATTAATCCAAAGCATTAATGAAACATCCCGAATTTCAATTAAGTTTGGG 501
Qy 228 SerCysThrAsnValAlaIleSerAsnLeuThrIleAsnAlaSerAlaLysSerProAsn 247
Db 502 AATTGTGAGGAGTAATAAATCATCGCATTAATGATTAAGGACCGAAGACATCTTAC 561
Qy 248 ThrAspGlyValHisIleSerAsnThrGlnIleGlnIleSerAspThrIleIleGly 267
Db 562 ACTGATGAAATGATATCTTTCATCTTAAAACTTTCACTTACAAAGAACACAGATAGGA 621
Qy 268 ThrGlyAspAspCysIleSerIleValIleSerGlySerGlnAsnValGlnAlaThrAsnIle 287
Db 622 ACAGGGATGATCTGCTGCTAATGAGCAGAGGCTTCTTAATTTGTGATTGAGGATCG 681
Qy 288 ThrCysGlyProGlyHisIleSerIleGlyIleSerLeuGlySerGlyAsnSerGlnAla 307
Db 682 ATTTGCGGTCCAGCGCATGGAATTAATGATTAAGTCTTGGGAGGAAACTCTAGAGA 741
Qy 308 TyrValSerAsnValThrValAsnGlnAlaLysIleIleGlyAlaGlnAsnGlyValArg 327
Db 742 GAGGTTTCATACCTGACGTAATGAGGCTTAATTCATAGACACACAAATGATTAAGA 801
Qy 328 IleLeuThrThrPheGlnGlyIleSerGlyGlnAlaSerAsnIleLysPheLeuAsnValGlu 347
Db 802 ATCAAAACATGGAGGGGTGGTTCAGGCATGCAAGCATATTAATTTATGAGAATTTTAA 861
Qy 348 MetGlnAspValLysTyrProIleIleIleAspGlnAsnTyrCysAspArgValGluPro 367
Db 862 ATGATTAATTCGAGAAACCCATTAATTAATCAATCTTACTGACCTTCACTTCTGCT 921
Qy 368 CysIleGlnPheSerAlaValGlnValLysAsnValAlaTyrGlnAsnIleLysGly 387
Db 922 TGCCAAACACAGAGGTGGGTGCTAAATCCAAATGATGACATCAACAGACATACGTGGG 981
Qy 388 ThrSerAlaThrLysValAlaIleLysPheAspCysSerThrAsnAsnProCysGluGly 407
Db 982 ACATCAGCAACAGCAGCAATTCATTAAGTCAAGTACAGAGTATGCTTGCAGAAAT 1041
Qy 408 IleIleMetGlnAsnIleAsnLeuValIleLysSerGlyLysProSerGlnAlaThrCys 427
Db 1042 ATTAAGCTAAGTATTAATCTTTGAACTTACTCTCAGGAAA-----ATTGCTTCTGC 1095
Qy 428 LysAsnValHisPheAsnAsn-----AlaGlnHisValThrProHisCysThrSerLeu 445

```

```

Db 1096 CTTATATATTAATGCAATGATATTTTCAATGAGACAGCATCCGATGCAAGAAATTTA 1155
Qy 446 GluIleSer 448
Db 1156 AGTCCAAGT 1164

RESULT 10
US-08-463-213-1
; Sequence 1, Application US/08463213
; Patent No. 5759829
; GENERAL INFORMATION:
; APPLICANT: SHEWMAKER, C.
; APPLICANT: KRIDL, J.
; APPLICANT: HIATT, W.
; APPLICANT: KNAUF, V.
; TITLE OF INVENTION: ANTI-SENSE REGULATION OF GENE
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Calgene, Inc.
; STREET: 1920 Fifth Street
; CITY: Davis
; STATE: CA
; COUNTRY: USA
; ZIP: 95616
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 2.0 MB
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 7.1
; SOFTWARE: Microsoft Word 5.1 (a)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,213
; FILING DATE: 5-JUNE-95
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/750,505
; FILING DATE: 27-AUG-91
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/240,408
; FILING DATE: 30-AUG-88
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 06/920,574
; FILING DATE: 17-OCT-86
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 06/845,676
; FILING DATE: 28-MAR-86
; ATTORNEY/AGENT INFORMATION:
; NAME: Elizabeth Laesen
; REGISTRATION NUMBER: 31,845
; NAME: Donna E. Scherer
; REGISTRATION NUMBER: 34,719
; NAME: Carl J. Schwedler
; REGISTRATION NUMBER: 36,924
; TELEPHONE/DOCKET NUMBER: CGNE 26-4
; TELEPHONE: (916) 753-6313
; TELEFAX: (916) 753-1510
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 219 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA to mRNA
; US-08-463-213-1

Alignment Scores:
Pred. No.: 2,15e-34 Length: 219
Score: 366.00 Matches: 70
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.59% Mismatches: 0
Query Match: 15.31% Indels: 0

```



```
DB: 1 Gaps: 0
US-10-691-374-2 (1-457) x US-08-463-213-1 (1-219)
QY 71 AengilylelyValleleAnValleuserPhegiyAlalyeglyAapGlylystThyTyr 90
   : : : : :
DB 7 CATGGGATTAAAGTGAATTAACTTACTTGGAGCTTAAGGCTGATGAAAAACATAT 66
QY 91 AepAenIIeAlaPhegluGlnAlaTTPaengIuaIaCyseSerSerArgThrProValGln 110
   : : : : :
DB 67 GATAATATATGCAATTCAGACAGCATGGAATGAAGCATGTTCACTTGAACACCTGTTCAA 126
QY 111 PheValIaIProlysaenlysaenlyrleuleuLyGlnIleThrPheSerGlyProCys 130
   : : : : :
DB 127 TTGTGGTTCCTAAACAAAGAAATTATCTTCACAAACAAATCACCTTTCAAGTCCATCC 186
QY 131 ArgSerSerIleSerVallysllePhegiySer 141
   : : : : :
DB 187 AGATCTTCTATTTCAGTAAAGATTTTGGATCC 219

RESULT 11
5453566-1
; Patent No. 5453566
; APPLICANT: SHEWMAKER, CHRISTINE K.; KRIDL, JEAN C.; HIATT,
; WILLIAM R.; KNAUF, VIC
; TITLE OF INVENTION: ANTISENSE REGULATION OF GENE EXPRESSION
; IN PLANT/CELLS
; NUMBER OF SEQUENCES: 2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/750,505
; FILING DATE: 27-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 240,408
; FILING DATE: 30-AUG-1988
; APPLICATION NUMBER: 920,574
; FILING DATE: 17-OCT-1986
; APPLICATION NUMBER: 845,676
; FILING DATE: 28-MAR-1986
; SEQ ID NO:1:
; LENGTH: 219
5453566-1

Alignment Scores:
Pred. No.: 2,15e-34 Length: 219
Score: 366.00 Matches: 70
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.59% Mismatches: 0
Query Match: 15.31% Indels: 0
Gaps: 0

US-10-691-374-2 (1-457) x 5453566-1 (1-219)
QY 71 AengilylelyValleleAnValleuserPhegiyAlalyeglyAapGlylystThyTyr 90
   : : : : :
DB 7 CATGGGATTAAAGTGAATTAACTTACTTGGAGCTTAAGGCTGATGAAAAACATAT 66
QY 91 AepAenIIeAlaPhegluGlnAlaTTPaengIuaIaCyseSerSerArgThrProValGln 110
   : : : : :
DB 67 GATAATATATGCAATTCAGACAGCATGGAATGAAGCATGTTCACTTGAACACCTGTTCAA 126
QY 111 PheValIaIProlysaenlysaenlyrleuleuLyGlnIleThrPheSerGlyProCys 130
   : : : : :
DB 127 TTGTGGTTCCTAAACAAAGAAATTATCTTCACAAACAAATCACCTTTCAAGTCCATCC 186
QY 131 ArgSerSerIleSerVallysllePhegiySer 141
   : : : : :
DB 187 AGATCTTCTATTTCAGTAAAGATTTTGGATCC 219

RESULT 12
5453566-1
; Patent No. 5453566
; APPLICANT: SHEWMAKER, CHRISTINE K.; KRIDL, JEAN C.; HIATT,
; WILLIAM R.; KNAUF, VIC
```

```
; TITLE OF INVENTION: ANTISENSE REGULATION OF GENE EXPRESSION
; IN PLANT/CELLS
; NUMBER OF SEQUENCES: 2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/750,505
; FILING DATE: 27-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 240,408
; FILING DATE: 30-AUG-1988
; APPLICATION NUMBER: 920,574
; FILING DATE: 17-OCT-1986
; APPLICATION NUMBER: 845,676
; FILING DATE: 28-MAR-1986
; SEQ ID NO:1:
; LENGTH: 219
5453566-1

Alignment Scores:
Pred. No.: 2,15e-34 Length: 219
Score: 366.00 Matches: 70
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.59% Mismatches: 0
Query Match: 15.31% Indels: 0
Gaps: 0

US-10-691-374-2 (1-457) x 5453566-1 (1-219)
QY 71 AengilylelyValleleAnValleuserPhegiyAlalyeglyAapGlylystThyTyr 90
   : : : : :
DB 7 CATGGGATTAAAGTGAATTAACTTACTTGGAGCTTAAGGCTGATGAAAAACATAT 66
QY 91 AepAenIIeAlaPhegluGlnAlaTTPaengIuaIaCyseSerSerArgThrProValGln 110
   : : : : :
DB 67 GATAATATATGCAATTCAGACAGCATGGAATGAAGCATGTTCACTTGAACACCTGTTCAA 126
QY 111 PheValIaIProlysaenlysaenlyrleuleuLyGlnIleThrPheSerGlyProCys 130
   : : : : :
DB 127 TTGTGGTTCCTAAACAAAGAAATTATCTTCACAAACAAATCACCTTTCAAGTCCATCC 186
QY 131 ArgSerSerIleSerVallysllePhegiySer 141
   : : : : :
DB 187 AGATCTTCTATTTCAGTAAAGATTTTGGATCC 219

RESULT 13
US-09-107-532A-2998
; Sequence 2998, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
```

NAME: Arinello, Pamela Deneke
 REGISTRATION NUMBER: 40,489
 REFERENCE/DOCKET NUMBER: GTC-012
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (781) 893-5007
 TELEFAX: (781) 893-8277
 INFORMATION FOR SEQ ID NO: 2998:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1356 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 MOLECULE TYPE: DNA (genomic)
 HYPOTHEICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: Enterococcus faecium
 FEATURE:
 NAME/KEY: misc. feature
 LOCATION: (B) LOCATION 1...1356
 SEQUENCE DESCRIPTION: SEQ ID NO: 2998:
 US-09-107-532A-2998

Alignment Scores:
 Pred. No.: 4,06e-25 Length: 1356
 Score: 299.00 Matches: 100
 Percent Similarity: 45.98% Conservative: 66
 Best Local Similarity: 27.70% Mismatches: 134
 Query Match: 12.51% Indels: 62
 DB: 4 Gaps: 15

US-10-691-374-2 (1-457) x US-09-107-532A-2998 (1-1356)

OY 73 l l e l y s v a l l i e a n v a l l e u s e r p h e g l y a l a l y s g l y a s p g l y l y s t h r t y r a s p a n 92
 DB 13 a t t a a t a t g a t g a t a t t t g a c a t t t g g c g c a t c a t t g a t ----- g a a t g a a t a c t 66
 OY 93 l l e l a p h e g l u g l i n a l t r p a n g l u a l c y s s e s e r a r g t h r p r o v a l g i n p h e a l 112
 DB 67 g a n c a a t a c a a c a a g c a t t g a t g c t g c a g c a a g c a g c g a g c a c c --- g r a g t t 123
 OY 113 v a l p r o l y s a e n l y s a e n l y t l e u l e u l y g l i n l e t h r p h e s e r g l y p r o c y s a a g s e r 132
 DB 124 g t t c a c a g a g t g a a --- t t t c t a c a g a g c t t a t t t t a a a a c a a t g t c g a a c t 179
 OY 133 s e r i l e s e r v a l l y s l e p h e g l y s e r l e u g l u a l a s e r s e r l y s i l e s e r a s p t y r l y 152
 DB 180 t c a t t t a a g c g a g a --- g c g g t t c g a a a t t c t g a c a t c c a a a a g a t t a t t c c 233
 OY 152 s ----- a s p a r g a r g l e u t r p ----- l l e a l a p h 160
 DB 234 t g t t g t t c a t t c a g t t g g a n a g a g t c c a t c g a a a a g t t a t g c t t c c t g a t c t a t g c 293
 OY 160 e a p s e r v a l g i n a n l e u v a l a g l y g l y t h r i l e a n g l y a n g l y g l n v a 180
 DB 294 t c a a a a t g r a g a a c a t c t g t c a c t g t t c g t a c a t a g a t g a a n t g a a a a a 353
 OY 180 l t r p t r ----- p r o s e r s e r c y s i s l i e a n l y s s e r l e u p r o c y 194
 DB 354 a r g t g c a t a c t t t a g a a t a g a a c a a t ----- a a t t c g c t t a t c c t --- 399
 OY 194 s a r g a s p a l a p r o t h r a l a l e u t h r p h e t r p a n c y s l y s a n l e u l y s v a l a e a n l e 214
 DB 400 ----- c g t c c t a a t g a t g a t g a g t t c c a t r a a c t g a t c a t g a n t t a c t g a a a g a t a t 452
 OY 214 u l y s e r l y s a e n l a g l i n l i l e h i s l e l y s p h e g l u s e r c y t h r a n v a l a l 234
 DB 453 t a a a t t g a t c c a g t a c c a a g t t g a c a g a c c a a c c a a t t c t t g c a g a t a t g c a a c t 512
 OY 234 a s e r l e u w e t l i e a n a l a s e r a l a l y s e r p r o a n t h r a s p g l y a l h s v a l s e 254
 DB 513 t g a t a a c t t g a c a t t t g a a c c a g c g a n t t a c c a a a t a t c a g a c g a a t t a t c c a a 572

OY 254 r a s n t h r g l n t y r i l e g l i n l e s e r a s p t h r i l e i l e g l y t h r g l y a s p a s c y s i l s e 274
 DB 573 a t c a t g c a a a a a t g a t g c g a t t a t g a t a t t g a c c a t a t t g a c g t g a g a t g a t c t a t t g c 632
 OY 274 r i l e a l s e r g l y s e r ----- g i n a n v a l g i n a l t h 285
 DB 633 a a t c a a g c a g a g a c t g a a a t a c g a a a g a t c c g t t c g a a a a a t a t c a g a t c a c 692
 OY 285 r a s n l e t h r c y s g l y p r o g l y h i s --- g l y l l e s e r l e g l y s e r l e u g l y s e r g l y a s 304
 DB 693 a a a t t c a c a g a t g c t c a t g a c a t g c g c g t g c t a t t a g a a g t a a a g a t g c g t a g 752
 OY 304 n s e r g l u a l a t y r v a l s e r a n v a l t h r v a l a n g l u a l a l y s l e i l e g l y a l a g l u a s 324
 DB 753 t ----- a t t c g a a a t a t t a c a t t t c a a t t g a t t c t t c c a a a a c t g a t c g 800
 OY 324 n g l y a l a g l i e l y t h r t r p g l n g l y s e r g l y g l n a l a s e r a n l e l y s p h e l e 344
 DB 801 a g a a t a c a t t g a a a t c t a g a c g a t c a g c g a g c g g a t c g t t g a a a t a t t c a g a t a t c 860
 OY 344 u a n v a l g l u e t g l n a s p v a l y s t y r p r o l i l e i l e a s p g l n a s n t y r --- c y s a s 363
 DB 861 c a a t a t a t c a t g a t a t a t g a t g a t g c t c a t t t a t a c t a c g a a t c t t a t t a t t t c g t g 920
 OY 363 p ----- a r g v a l g l u p r o c y s i l e g l i n l ----- 371
 DB 921 t c c g c g a g a a a a a c c g t a t g a t g g a a a a a a c g t t a t c c a t a g a t a a c g a c 980
 OY 372 ----- p h e s e r l a v a l g i n v a l l y s a n v a l a t y r g l u a n l e l y s g l y t h r s e 389
 DB 981 a c c a g c a t t c a g a a g a t c a t t t t t c a a t a t c a c g c a a g a a a t g c a t c t t c t c g 1040
 OY 389 r 389
 DB 1041 A 1041

RESULT 14
 US-09-614-221A-573
 Sequence 573, Application US/09614221A
 Patent No. 6723837
 GENERAL INFORMATION:
 APPLICANT: Karunanandaa, Balasubramanian
 APPLICANT: Yu, Jaehyuk
 APPLICANT: Kishore, Ganesh M.
 TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED
 FILE REFERENCE: 16516.075
 CURRENT APPLICATION NUMBER: US/09/614,221A
 PRIOR FILING DATE: 2000-07-12
 PRIOR APPLICATION NUMBER: US 60/142,981
 NUMBER OF SEQ ID NOS: 626
 SEQ ID NO 573
 LENGTH: 1086
 TYPE: DNA
 ORGANISM: Saccharomyces cerevisiae
 US-09-614-221A-573

Alignment Scores:
 Pred. No.: 6,37e-25 Length: 1086
 Score: 296.00 Matches: 112
 Percent Similarity: 44.85% Conservative: 58
 Best Local Similarity: 29.55% Mismatches: 147
 Query Match: 12.38% Indels: 62
 DB: 4 Gaps: 17

US-10-691-374-2 (1-457) x US-09-614-221A-573 (1-1086)

OY 61 s e r a n s a n a n l i e a p l y s a n g l y l e l y s v a l l i e a n v a l l e u s e r 80
 DB 97 t c t t t g c t c a c t c t a a c c g t g a a a a a t t r a n g c a g c a t c g t a t a t a a g a c t t a a c t 156
 OY 81 p h e g l y a l a l y s g l y a s p g l y l y s t h r t y r a s p a n l e a l a p h e g l u g l i n a l t r p a n 100

Db	157	GTCCACGCT-----				GGACAGACCTTAAAGTTTAACCTGGGTTA-----	192
Oy	101	GIUALaCySerSerArgThrProValGlnPheValProLysAsnLysAsnTyrLeu					120
Db	193	-----AGCAGTGGTACTACTGTGTACGTTTGAAGGACAAACCACTTCAGTACGAG					243
Oy	121	LeuLysGlnIleThrPheSerGlyProCysArgSerSerIleSerValLysIlePheGly					140
Db	244	-----GAATGGAGCGGCCCTTTA-----ATTTCATC-----					270
Oy	141	SetLeuGlnAlaSerSerLysIleSerAspTyrLysAspArgArgLeuTrpIleAlaPhe					160
Db	271	-----TCAGGGCTTAATAATCAGC-----					288
Oy	161	AepSerValGlnAsnLeuValAlaGlyIleGlyGly---ThrIleAsnGlyAsnGlyGln					179
Db	289	-----CTTGTTGGTCTGTCGGGACATACCATTCATATGCTCAGAGACA					330
Oy	180	ValTyrTrpProSerSerSerCysLysIleAsnLysSerLeuProCysArgIleAlaProThr					199
Db	331	AAATGGTGGAGTGGCTTAGTGATGACGGGTAAAGTCAACCG---AAGTTGTAAAGTTG					387
Oy	200	AlaLeuTrpThrPheTrpAsnCysLysAsnLeuValAlaAsnLeuLysSerLysAsnAla					219
Db	368	GGGTGTAGAG-----GGAACATCTAAGGTCAACCGATGTGAATTTAAATAATGCT					435
Oy	220	GlnGlnIleHisIleLysPheGlySerCysTrpAsnValAlaIleSerAsnLeuMetIle					239
Db	436	CCACACCAAGTTTCCAGCATCAATAATGTTTCAGATTTAACCATCAGCACATACAAATT					495
Oy	240	AsnAla-----SerAlaLysSerProAsnThrAspGlyValHisValSerAsn					255
Db	496	GAATACAGAGACGGTATTCGGCTGGTGCTCATTAATACGAGAGGGTTGATGTGGTACT					555
Oy	256	ThrGlnTyrIleGlnIleSerAspThrIleIleGlyThrGlyAspAspCysIleSerIle					275
Db	556	TCTACTAAAGCTTTAATTCAGAGATGATCTGTTTAAATCAGAGATGATCTGATGCTGCG					615
Oy	276	ValSerGlySerGlnAsnValGlnAlaThrAsnIleThrCysGlyProGlnIleGlyIle					295
Db	616	AAATCCGGTTCA---ACTATTAAATTTATGAACCACTACTGTCACATGGCCATGGTATT					672
Oy	296	SerIleGlySerLeuGlySerGlyAsnSerGlnAlaTyrValSerAsnValThrValAsn					315
Db	673	TCTGTAAGTTCGTTGGT---GGCCGTTCTGATTAATACAGTCATAGTTCGGGCTGAA					729
Oy	316	GluAlaLysIleIleGlyAlaGluLysGlyValAlaGlyIleLysThrTrpGlnGlyGlySer					335
Db	730	AATAACCAATGTTATCAACTGACAAACGGGTGAGAAATAAAAACCGTAGAAGGTGGACA					789
Oy	336	GlyGlnAlaSerAsnIleLysPheLeuAsnValGluMetGlnAspValLys---TyrPro					354
Db	790	GGCACAGTCATCAATGCTCACTTATACGTAATTAATAATTAAGCGGCATTAAGTTATGGT					849
Oy	355	IleIleIleAspGlnAsnTyrCysAspArgValGluProCysIleGlnGlnPheSerAla					374
Db	850	ATTGTTATCGAAGGCGATTATTGGAATACT---AGACATACCGAACTGCTACAGTGGC					906
Oy	375	ValGlnValLysAsnValValTyrGluAsnIleLysGly-----ThrSerAlaThr					391
Db	907	GTTCCCAATTCGAATTTAGTATGAAGATATCAACGGGAGGCGTGAATCCACAGGAGAG					966
Oy	392	LysValAlaIleLysPheAspCysSerThrAsnPheProCysGluGlyIle---IleMet					410
Db	967	AGGGTTAAATTTGGTGAATAACGCTACTTAATCGGCATATGCTGGGGGTGCAATTACC					1022
Oy	411	GluAsnIleAsnLeuValGlyLysGlyLysPro-----SerGluAlaThrCys					427
Db	1027	GGTGGTCTCTCTATTCGAGAGTTCTGGAATCCCATCGAGATTCGGTTCGAACGCTGT					1083

```

; Sequence 1, Application US/09787583
; Patent No. 6602696
; GENERAL INFORMATION:
; APPLICANT: DSM N.V.
; APPLICANT: France, Maartje
; APPLICANT: Grassin, Catherine
; APPLICANT: Herreijer, Margareta
; APPLICANT: Meuwesen, Petrus
; APPLICANT: Ooijen, Albert
; APPLICANT: Voragen, Alphons
; TITLE OF INVENTION: ASPERGILLUS TUBIGENSIS POLYGALACTURONASE
; FILE REFERENCE: 24615-2014.00
; CURRENT APPLICATION NUMBER: US/09/787,583
; CURRENT FILING DATE: 2001-03-19
; PRIOR APPLICATION NUMBER: EP 99200481.2
; PRIOR FILING DATE: 1999-02-17
; PRIOR APPLICATION NUMBER: EP 98203171.8
; PRIOR FILING DATE: 1998-09-18
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1570
; TYPE: DNA
; ORGANISM: Aspergillus tubingensis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (172) ... (1260)
US-09-787-583-1

Alignment Scores:
Score: 2,33e-24 Length: 1570
Pct Ident: 293.50 Matches: 81
Percent Similarity: 50.54% Conservative: 59
Best Local Similarity: 29.24% Mismatches: 96
Query Match: 12.28% Indels: 41
DB: Gaps: 14

US-10-691-374-2 (1-457) x US-09-787-583-1 (1-1570)

Oy 171 GlyGlyThrIleAsnGlyValInValTTP----- 182
Db 475 GTGCCAAGATCACTCGACGGTGCTGCCTGTGGAGCAGCGGATGACCAAGGTGC 53
Oy 183 -----ProSerSetCyelysIleAsnLysSerLeuProCySarghapAlaProthr 199
Db 535 AAGACCAAGCCCAAGTTCTTCAGGTCCAAG----- 566
Oy 200 AlaIeuThrPheTrpAsnCyeLyAsnLeuLysValAsnAsnLeuLysSerLyAsnAla 219
Db 568 ---CTCGACGAGTCCAGCATCACCGGCTCGAAGATCTCAACAACCCCCTGCCAGGCTTC 624
Oy 220 GluGlnIle-----HisIleLysPheGluSerCySThAsnValValAlaSerAsnLeu 237
Db 625 AGCATTTGGCTGACCACCTG-----ACCATCACTGACGTGAACATTGACCACTCC 675
Oy 238 MetIleAsnAlaSerAlaLysSerProAsnThrAspGlyValHisValSerAsnThrGln 257
Db 676 GCC-----GGACGACGACAGGCGCACAAACCGATGCTTTGACATTGGTCAAGATACC 729
Oy 258 TyrIleGlnIleSerAspThrIleIleGlyThrGlyAspAspCysIleSerIleValSer 277
Db 730 TACATCAACATCGACGATGCCACCGCTCAACACAGATGATGCTGCGCATCACTCG 769
Oy 278 GlySerGlnAsnValGlnAlaThrAsnIleMetCySgLyProGlyYhiGlyIleSerIle 297
Db 790 GGT---GGACACATCACTTTCACCAACGGTTACTGTGACGGTGGCCACGGTCTCTCCATT 846
Oy 298 GlySerLeuGlySerGlyAsnSerGlnAlaTyrValSerAsnValThrValAsnGlnAla 317
Db 847 GGTTCATTGTTGTT---GGCCGACGACACAAACCGTCAACGACGTGACATCTCAACTCC 903
Oy 318 LysIleIleIleGlyAlaGlnAsnGlyValArgIleLysThrTTPGlnGlyGlySerGlyGln 337

```

```
Db      904 AAGTGCTCACTCCGAAACGGTGTCCGATCAGACCATCTACGGCAAGACCGGCACT 963
QY      338 AlaSerAsn1LeIysPheLeuAsnValGluMetGlnAspVal--LysTyrProIleIle 356
Db      964 GTTGAGAACGTCMAAGTTTCGAGGACATCACCTGTCCGACATCAGCAAGTACGGTATCGTC 1023
QY      357 IleAspGlnAsnTyrCysAspArgValGluProCysIleGlnGlnPheSerAlaValGln 376
Db      1024 GTTGACGACGACTAC--GAGAACGGCAGCCGCCACCGGCACGCCACCAACGAGTGTCAAG 1080
QY      377 ValLysAsnValValTyrGlnAsn1LeIysGly-----ThrSerAlaThrLysVal 393
Db      1081 GTTGAGGACATCATCTTCAAGAGGTCAACCGGACCGTCACAGAGCTGTGTACTGACATC 1140
QY      394 AlaIleLysPheAspCysSerThrAsnPheProCysGluGlyIleIleMetGluAsnIle 413
Db      1141 TACATCTCTG-----TGGGTTCCGGC--AGCTGCTCGAACTGGACCTGGAGCGGTGTT 1191
QY      414 AsnLeuValGlyGluSerGlyLysProSerGluAlaThrCysLysAsnVal 430
Db      1192 GATGTGACCGGC-----GGCAAGAAAGAC--AGCAAGTGCAGAAAGACGTC 1233
```

Search completed: March 5, 2005, 02:29:24
Job time : 251 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 4, 2005, 12:45:39 ; Search time 640 Seconds
(without alignments)
4227.065 Million cell updates/sec

Title: US-10-691-374-2

Perfect score: 2390

Sequence: 1 MTIGNSILLIIIFASIS.....VTPHCTSLISDEALYNY 457

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlp
-Q/cgn2.1/USPTO.epool_p/US10691374/tunat.28022005.120705.20976/app_query.fasta_1.647
-DB=N Geneseq.16Dec04 -QFMT=fastap -SUFFIX=mg -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNIT=bits -START=1 -END=-1 -MATRIX=blomsum62 -TRANS=human0.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pclo -NOR=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10691374_@CGN_1_1_708_@tunat.28022005.120705.20976 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: N_Geneseq_16Dec04:*
2: geneseqn1980s:*
3: geneseqn1990s:*
4: geneseqn2000s:*
5: geneseqn2001as:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2390	100.0	1624	1	AAN91112
2	2390	100.0	1636	2	AAQ38415
3	2390	100.0	1636	4	AAC84653
4	2390	100.0	1637	1	AAN80487
5	1820	76.2	5822	6	AAS20853

6	1127	47.2	1296	6	ADG87885
7	1127	47.2	1296	8	ADA68531
8	1127	47.2	1296	3	AA646786
9	1115.5	46.7	1280	8	ADA68538
10	1113	46.6	1631	2	AA63603
11	1103	46.2	1657	2	AA633994
12	1103	46.2	1657	2	AA222980
13	1022.5	42.8	1659	8	ADA70539
14	1011	42.3	1182	8	ADA70540
15	999	41.8	1144	8	ADA70633
16	971.5	40.6	1168	12	ADJ39618
17	925	38.7	1603	4	AA003721
18	856.5	35.8	4783	2	AA056765
19	845	35.4	1505	3	AA050895
20	841.5	35.2	1236	8	ADA69348
21	829	34.7	5002	2	AA255011
22	791.5	33.1	1772	2	AA638521
23	787.5	32.9	1545	2	AA090156
24	787	32.9	1542	2	AA084045
25	787	32.9	1726	2	AA066048
26	787	32.9	1733	2	AA084046
27	787	32.9	1733	2	AA618102
28	786.5	32.9	1479	2	AA066051
29	768	32.1	1395	2	AA066050
30	768	32.1	1410	2	AA066049
31	767	32.1	1380	2	AA084044
32	760	31.8	1524	4	AA66167
33	760	31.8	1759	4	AA66166
34	727	30.7	1512	8	ADA70810
35	727	30.4	1673	6	AA032961
36	718	30.0	1492	3	AA032994
37	709.5	29.7	1501	3	AA046834
38	702.5	29.4	1185	6	AB214715
39	694	29.1	1155	8	ADA67805
40	694	29.0	1496	2	AA056767
41	690	28.9	1736	3	AA042340
42	676.5	28.3	1531	3	AA049688
43	673	28.2	1590	3	AA039627
44	658.5	27.6	1059	3	AA036631
45	652	27.3	1182	8	ADA70913

ALIGNMENTS

RESULT 1	
ID	AAN91112 standard; DNA; 1624 BP.
XX	
AC	AAN91112;
XX	
DT	25-MAR-2003 (revised)
DT	21-JUN-1990 (first entry)
XX	
DE	Polysaccharuronase (PG) gene cDNA.
XX	
KW	Polysaccharuronase gene; tomato; ds.
XX	
OS	Lycopersicon esculentum.
XX	
FT	Key
FT	CDS
FT	Location/Qualifiers
XX	47..1418
XX	/tag= a
XX	
PN	US4801540-A.
XX	
PD	31-JAN-1989.
XX	
PF	02-JAN-1987; 87US-00000201.
XX	
PR	28-MAR-1986; 86US-00845676.
PR	17-OCT-1986; 86US-00920574.
XX	
PA	(CALJ) CALGENE INC.

Adg87885	A. thalia.
Ada68531	Arabidops
Aac6786	Arabidops
Ada68538	Arabidops
Aac63603	Dehiscenc
Aac33994	Oilseed r
Aa222980	Brasica
Ada70539	Rice gene
Ada70540	Rice gene
Ada70633	Rice gene
Adj39618	Plant cDN
Aad03721	Tomato po
Aag6765	Putative
Aac50895	Arabidops
Ada69348	Rice gene
Aax25011	Anther-ep
Aat38521	Chamaecyp
Aag90156	Japanese
Aag84045	Japanicum
Aag66048	Japanese
Aag84046	Japanicum
Aat18102	Japan ced
Aag66051	Japanese
Aag66050	Japanese
Aag66049	Japanese
Aac84044	Japanicum
Aat66167	Uniperus
Aat66166	Uniperus
Ada70810	Rice gene
Aad32961	Pear poly
Aac32994	Arabidops
Aac6834	Arabidops
Ab214715	Arabidops
Ada67805	Arabidops
Aac6767	Putative
Aac42340	Arabidops
Aac9688	Arabidops
Aac39627	Arabidops
Aac36631	Arabidops
Ada70913	Rice gene

XX Hiact WR, Sheehy RE, Shemmaker CK, Kridl JC, Knauf V,
 XX WPI: 1989-053640/07.
 DR P-PSDB; AAP94619.
 XX
 PT Tomato polygalacturonase gene - used for modulating expression in plant
 cells or directing expression of heterologous peptide(s).
 XX
 PS Disclosure; Page ?; bpp; English.
 XX
 CC cDNA of gene may act as a probe to the genomic sequence, provides a means
 CC of modulating the production of PG and acts as a source of the transic
 CC peptide which may be joined to heterologous peptides directing them to
 CC cell wall. (Updated on 25-MAR-2003 to correct PA field.)
 XX

SQ Sequence 1624 BP; 586 A; 231 C; 294 G; 513 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	6.03e-219	Length:	1624
Score:	2390.00	Matches:	457
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	1	Gaps:	0

US-10-691-374-2 (1-457) x AAN91112 (1-1624)

QY 1 MetValIleGlnArgAsnSerIleuLeuLeuIleIleIlePheAlaSerSerIleSer 20
 Db 47 ATGGTATCCAAAGAAATGATATCTCTCTCATATATATTTGGTTCATCAATTTCA 106
 QY 21 ThrCysAspSerAsnValIleAspAspAsnLeuPheLysGlnValTyrAspAsnIleLeu 40
 Db 107 ACTGTGACAAACATGTTATGATGATGATGATGATGATGATGATGATGATGATGAT 166
 QY 41 GlnGlnGlnPheAlaHisAspPheGlnAlaTyrLeuSerTyrLeuSerTyrLeuSer 60
 Db 167 GAACAAATTTGCTCATGATTTTCAAGCTTATCTTCTTATTTGAGCAAAATATTTGAA 226
 QY 61 SerAsnAsnAsnIleAspLysValAspLysValIleLysValIleAsnValIleAsn 80
 Db 227 AGCAACATATATATTTGCAAGTTGATTAATGATTAATGATTAATGATTAATGATTA 286
 QY 81 PheGlyAlaLysGlyAspGlyLysGlyTyrAspAsnIleAlaPheGlnGlnAlaTyrAsn 100
 Db 287 TTGGAGCTAAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 346
 QY 101 GlnAlaCysSerSerArgThrProValGlnPheValValProLysAsnLysAsnTyrLeu 120
 Db 347 GAAGCATGTTCACTGAAACACCTGTTCAATTTTGCTTCTTAAACAAAGAAATATCTT 406
 QY 121 LeuLysGlnIleThrPheSerGlyProCysArgSerSerIleSerValLysIlePheGly 140
 Db 407 CTCACCAATATCACTTTTCAGGTCATCAGATCTTCAATTTCAAGTAAGATTTTGA 466
 QY 141 SerLeuGlnAlaSerSerLysIleSerAspTyrLysAspArgLysLeuTyrIleAlaPhe 160
 Db 467 TCTTTGAAGCATCTAGTAAATTTCAAGCTAACAAAGATGAAAGCTTTGATGCTTTT 526
 QY 161 AspSerValGlnAsnLeuValValGlyGlyGlyTyrIleAsnGlyAsnGlyGlnVal 180
 Db 527 GATAGGTCTCAAAATTTAGTTGTGAGAGAGAGAACTATCAATGCAATGCAATGCA 586
 QY 181 TrpTrpProSerSerCysLysIleAsnLysSerLeuProCysArgAspAlaProThrAla 200
 Db 587 TGTGTCCCAAGTTTCATGCAAAATTAATAATCACTGCCATCAGGAGATGACCAACGCC 646
 QY 201 LeuThrPheTrpAsnCysLysAsnLeuLysValAsnAsnLeuLysSerLysAsnAlaGln 220
 Db 647 TTAACTCTTGGAATTTGCAAAATTTGAAAGTGAATTAATCTTAAAGAGTAAAGATGCA 706
 QY 221 GlnIleHisIleLysPheGlnLysCysThrAsnValValAlaSerAsnLeuMetIleAsn 240

Db 707 CAATTATATCAAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 766
 QY 241 AlAserAlaLysSerProAsnThrAspGlyValIleValSerAsnThrGlnTyrIleGln 260
 Db 767 GCTTCAGCAAAAGGCCCAATCTGATGAGTCCATGATCAAAATATCAATATATTTCA 826
 QY 261 IleSerAspThrIleIleGlyThrGlyAspAspCysIleSerIleValIleSerGln 280
 Db 827 ATATCTGATCTATTTTGAACAGGTATGATGATGATGATGATGATGATGATGATGAT 886
 QY 281 AsnValGlnAlaThrAsnIleThrCysGlyProGlyIleGlyIleSerIleGlySerLeu 300
 Db 887 AATGTCAGAGCCCAAAATATTTACTTGTGTCAGGTATGATGATGATGATGATGATGAT 946
 QY 301 GlySerGlyAsnSerGlnAlaTyrValSerAsnValThrValAsnGlnAlaLysIleIle 320
 Db 947 GATCTGCAAAATTCAGAAAGCTTATGTCCTAATGTTACTGTAAATGAAAGCCAAATATATC 1006
 QY 321 GlyAlaGlyAsnGlyValArgIleLysThrTrpGlnGlyGlySerGlyGlnAlaSerAsn 340
 Db 1007 GGTGCGCAAAATGAGATTAGATCAAGCTTGCAAGGAGATCTGCAAGCTGATGCAAC 1066
 QY 341 IleLysPheLeuAsnValGlnMetGlnAspValLysTyrProIleIleIleAspGlnAsn 360
 Db 1067 ATCAAAATTTCTGAATGTGAAATGCAAGACCTTAAGATATCCATTAATTATAGACCAAAAC 1126
 QY 361 TyrCysAspArgValGlnProCysIleGlnGlnPheSerAlaValGlnValLysAsnVal 380
 Db 1127 TATGTGATCGAGTTGAACCATGTRATCAACAGTTTTCAGAGTTCAAGGAAATGATG 1186
 QY 381 ValTyrGlnAsnIleLysGlyThrSerAlaThrLysValAlaIleLysPheAspCysSer 400
 Db 1187 GTGATGGAATATCAAGGCGCAAGTCAAAAGGTGCCATTAATTTGATTTGAGC 1246
 QY 401 ThrAsnPheProCysGlnGlyIleIleMetGlnAsnIleAsnLeuValGlyLysGly 420
 Db 1247 ACAAACTTTCATGATGAAGAAATTAATATGAGAAATTAATTAATTAATTAATTAATTA 1306
 QY 421 LysProSerGlnAlaThrCysLysAsnValHisPheAsnAsnAlaGlnHisAlaTyrPro 440
 Db 1307 AAACCATCAAGAGGCTACGTCAAAATATGTCATTTTAACATCTGAACATGTTAACCA 1366
 QY 441 HisCysThrSerLeuGlnIleSerGlnAspGlnAlaLeuLeuTyrAsnTyr 457
 Db 1367 CACTGCCTTCACTAGAAATTTCAAGAGATGAAAGCTCTTTGTATATATAT 1417

RESULT 2
 AAQ38415
 ID AAQ38415 standard; DNA; 1636 BP.
 XX
 AC AAQ38415;
 XX
 DT 27-AUG-2003 (revised)
 DT 25-MAR-2003 (revised)
 DT 05-JUL-1993 (first entry)
 XX
 DE Polygalacturonase cDNA clone pTOM6.
 XX
 KW polygalacturonase; PG; pectin esterase; PE; expression regulation;
 KW fruit softening enzymes; flowering plants; fruiting plants;
 KW antisense RNA.
 XX
 OS Lycopersicon esculentum.
 XX
 FH Key Location/Qualifiers
 FT CDS 51..1422
 FT /*tag= a
 FT /label= polygalacturonase (PG)
 XX
 PN EF532060-A1.
 XX
 PD 17-MAR-1993.

XX 06-NOV-1987; 92EP-00117411.
PP
XX
PR 11-NOV-1986; 86GB-00026879.
XX
PA (ICIL) IMPERIAL CHEM IND PLC.
PA (ZENE) ZENECA LTD.
XX
PI Bridges IG, Grieron D, Schuch WM;
XX
DR WPI; 1993-087084/11.
DR P-PSDB; AAR32107.
XX
PT Recombinant DNA for flowering and fruiting plants e.g. tomatoes ripening
PT control - comprises base sequence for transcription contg. inverted
PT sequence of bases complementary to bases in anti sense ribonucleic acid
XX encoding softening enzymes, or gene expression regulation.
XX
PS Example 12; Fig 1; 20pp; English.

XX This is the sequence of clone pTOM6 which contains the coding sequence of
CC polygalacturonase (PG). It was used to isolate the PG promoter in the
CC construction of a vector encoding antisense RNA to the PG cDNA and PG
CC gene. This would be useful to regulate the expression of the fruit
CC softening enzymes in flowering and fruiting plants. Such antisense RNA
CC would delay fruit softening. (Updated on 25-MAR-2003 to correct PN
CC field.) (Updated on 25-MAR-2003 to correct PF field.) (Updated on 25-MAR-
CC 2003 to correct PA field.) (Updated on 27-AUG-2003 to correct OS field.)
XX

SO Sequence 1636 BP; 598 A; 232 C; 290 G; 516 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 6, 09e-219 Length: 1636
Score: 2390.00 Matches: 457
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-10-691-374-2 (1-457) x AAC38415 (1-1636)

QY 1 MetValIIGLhArgnSenSerIleuLeuLeuIleIlePheAlaSerSerIleSer 20
DB 51 ATGCGTATCCAAAGAGATGATTCCTTCCTCATTAATTAATTTCTTCATCAATTTCA 110
QY 21 ThrCyArgSerSerValIleAAspAAsnLeuPheLeuGlnValTyAspAAsnIleLeu 40
DB 111 ACTGTAGAAAGCAATGTTATGATGACAAATTTATTCAAACAAGTTATGATATATCTT 170
QY 41 GlnGlnGluPheAlaHisAspPheGlnAlaTyLeuSerTyLeuSerTyLeuSerTyLeu 60
DB 171 GAACAAGAAATTCCTCATGATTTTCAAGCTTATCTTTATTTGAGCAAAATATTTGAA 230
QY 61 SerAAsnAAsnIleAspTyValAspTyValAspTyValIleuValIleAsnValIleuSer 80
DB 231 AGCAACAATAATATTAACAAGGTGATTAATAAATGGAATTAAGATTAATGACTTAC 290
QY 81 PheGlyAlaLeuGlyAspGlyTyLeuTyAspAAsnIleAlaPheGlnGlnAlaTyPAsn 100
DB 291 TTTCAGCTTAAGGCTATGGAATAACATATGATTAATTCATTGACGAAGCATGAAT 350
QY 101 GlnAlaCySerSerArgTyProValGlnPheValIleProTyAsnTyLeuSerValIlePheGly 120
DB 351 GAAGCATGTTCACTGAAACACCTGTTCAATTTGTGTTCTTAAACAAAGAAATTAATCTT 410
QY 121 LeuTyGlnIleThrPheSerGlyProCyArgSerSerIleSerValIlePheGly 140
DB 411 CTCAGCAAAATCACCTTTTCAGGTCCATGCAATCTTCTAATTCAGTAAGATTTTGG 470
QY 141 SerLeuGlnAlaSerSerTyValIleAspTyTyLeuTyAspArgTyLeuTyPAsnIleAlaPhe 160
DB 471 TCCTTAAGAACATCTAATAAATTTTCAGACTCAAAAGATTAAGAGGCTTTGGATTTGCTTTT 530

QY 161 AAspSerValGlnAAsnLeuValValGlyGlyGlyTyThrIleAAsnGlyAAsnGlyGlnVal 180
DB 531 GATAGTGTTCAAAATTTAGTTGTGAGAGAGAGAACTATCAATGCGCATATGACAAAGTA 590
QY 181 TrpTyPProSerSerCyLeuTyValIleAAsnTySerLeuProCyArgAAspAlaProThrAla 200
DB 591 TGTGTGCCAAGTCTTCGCAAAATTAATTAATCACTGCCATGCAAGGATGCAACACGGCC 650
QY 201 LeuThrPheTyPAsnCyValTyAsnLeuTyValAAsnAAsnLeuTySerTyAsnAlaGln 220
DB 651 TTAACCTTCTGGAAATTCGCAAAATTTGAAGGAAATTAATCTAAGAGTAAAAATGCACAA 710
QY 221 GlnIleHisIleLeuPheGlnTySerTyPAsnValIleAlaSerAAsnLeuMetIleAAsn 240
DB 711 CAAATTCATCAAAATTTGAGTCATGCACTAATGTTGCTTCAATTTGATGATCAAT 770
QY 241 AlaSerAlaTySerProAAsnThrAspGlyValHisValSerAAsnThrGlnTyIleGln 260
DB 771 GCTTACGAAAGAGCCCAAAATCTGATGAGTCCATGTAACAAATCTCAATATATTCA 830
QY 261 IleSerAspThrIleIleGlyTyThrGlyAspAAspGlyIleSerIleValSerGlySerGln 280
DB 831 ATATCTGATTAATTAATGAAACAGATGATGATTTGATTTCAATTTCTGATCTCA 890
QY 281 AAsnValGlnAlaThrAAsnIleThrCyGlyTyProGlyHisGlyIleSerIleGlySerLeu 300
DB 891 AATGTGCAAGGCCCAAAATTAATCTTGCTGCTCAAGTCAAGTCAATTAAGTAATGAAAGCTTA 950
QY 301 GlySerGlyAAsnSerGlyValTyValSerAAsnValThrValAAsnGlnAlaTyIleIle 320
DB 951 GATCTGCAAAATTCGAAAGCTTATGCTAATGATTAATGAAATGAAAGCCAAATTAATTC 1010
QY 321 GlyAlaGlnAAsnGlyValArgIleTyThrTyProGlnGlyGlySerGlyGlnAlaSerAAsn 340
DB 1011 GGTGCCGAAATGAGAGTATGATGATCAAGCTTGCGAGAGAGATCTGCAACACCTAGCAAC 1070
QY 341 IleTyPheLeuAAsnValGlnMetGlnAAspValTyTyProIleIleIleAAspGlnAAsn 360
DB 1071 ATCAAAATTTCTAATGAGAAATGCAAGACGTTAAGTATCCATTAATTAATTAATTAATTA 1130
QY 361 TyTyAspArgValGlnTyProCyValIleGlnIlePheSerAlaValGlnValTyAAsnVal 380
DB 1131 TATTTGATCGAGTGTGAACATGATATACAAAGTTTTCGCGATTCAGTCAATGAAATTAATG 1190
QY 381 ValTyTyGlnAAsnIleTyGlyTyThrSerAlaThrTyValAlaIleTyPheAAspCySer 400
DB 1191 GTGATGAGAAATATCAAGGCGCAAGTGCACAAAGTGCATTAATAATTTGATTCACGC 1250
QY 401 ThrAAsnPheProCyGlnGlyIleIleMetGlnAAsnIleAAsnValGlyGlySerGly 420
DB 1251 ACAAACTTTCATGTAAGAAATTAATTAATGAGAAATTAATTAATTAATTAATTAATTAATTA 1310
QY 421 TyPProSerGlyAlaThrCyValTyAAsnValHisPheAAsnAlaGlnHisValTyPro 440
DB 1311 AAACATCAAGAGCTACGTCGCAAAATGTCATTTTAACAAGTCGCAACATGTTATACCA 1370
QY 441 HisCyThrSerLeuGlnIleSerGlyAAspGlnAlaLeuLeuTyTyAsnTy 457
DB 1371 CACTGCACTTCACTAGAAATTTTCAGAGATGAAGCTTTTGTATTAATTAAT 1421

RESULT 3
AAC84653
ID AAC84653 standard; cDNA; 1636 BP.
XX
AC AAC84653;
XX
DT 20-APR-2001 (first entry)
XX
DE cDNA sequence of pTOM6 encoding a polygalacturonase (PG) enzyme.
XX
XX Pectin; pectin methylase; PMG; polygalacturonase; PG; tomato;
KM food product; yogurt; milk; fruit juice; whey drink; de-esterification;
KM pTOM6; ss.

XX OS Lycopersicon esculentum.
 XX FH Key Location/Qualifiers
 FT CDS 51..1424
 FT /*tag= a
 FT /product= "PG enzyme"
 PN MO200078982-A1.
 XX 28-DEC-2000.
 PD 15-JUN-2000; 2000WO-IB000869.
 XX 17-JUN-1999; 99GB-00014209.
 XX (DANI-) DANISCO AS.
 XX Christensen TME, Krelberg JD;
 PI WPI; 2001-091573/10.
 DR P-PSDB; AAB48338.
 XX
 XX Modifying pectin, for foodstuffs preparation, involves transforming host
 PT having pectin methyltransferase (PME) and polygalacturonase (PG) activity by
 PT silencing PG activity, to increase PME to PG ratio.
 XX
 XX Claim 6; Fig 1; 78pp; English.
 XX
 CC The invention provides a new method for modifying pectin that involves
 CC providing a host having pectin methyltransferase (PME) activity and
 CC polygalacturonase (PG) activity, transforming the host by silencing PG
 CC activity to provide an increased PME to PG ratio, preparing a PME extract
 CC from the transformed host, and using the PME extract to modify pectin. A
 CC PME modified pectin is useful for foodstuffs preparation, and to impart
 CC an increased functionality to food products such as yogurt, milk/fruit
 CC juice and whey drinks. PME is useful to reduce the number of ester groups
 CC in a pectin in a block-wise manner, and to de-esterify two or more
 CC adjacent galacturonic acid residue of a pectin on at least substantially
 CC all of the pectin chains. The present sequence represents the cDNA
 CC sequence of pTOM6 encoding a PG enzyme. A nucleotide sequence (seq id No.
 CC 3) encoding the PG enzyme which is deposited as pTOM23 with NCIM
 CC (Accession number 12373) is also claimed. The sequence for seq id No. 3
 CC is not provided in the specification
 SQ Sequence 1636 BP; 598 A; 232 C; 290 G; 516 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 6.09e-219 Length: 1636
 Score: 2390.00 Matches: 457
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 4 Gaps: 0
 US-10-691-374-2 (1-457) x AAC84653 (1-1636)
 QY 1 MetValIIGlAArgAsnSerIleuLeuIleIleIlePheAlaSerSerIleSer 20
 DB 51 ARGGGTATTCAGAGAAATGATTCCTTCATTAATTAATTTGGCTTCATCAATTTCA 110
 QY 21 ThrCysArgSerAsnValIleAspAspAsnLeuPheIleValIleAspAsnIleLeu 40
 DB 111 ACTTGTAGAACGAAAGTATTATGATGCAATTTATTCAAACAGTTTATGATTAATTCCT 170
 QY 41 GluGlnGluPheAlaHisAspPheGlnAlaTyrIleuSerTyrIleuSerTyrAsnIleGlu 60
 DB 171 GAACAAGAAATTTGCTCATGATTTTCACCTTATCTTTATTGAGCAAAAATATTGAA 230
 QY 61 SerAsnAsnAsnIleAspIleValAspIleAsnGlyIleIleValIleAsnValIleuSer 80
 DB 231 AGCAACATTAATATTGACAGAGTTGATTAATAATGAGATTAAAGATTAAATGACTTAC 290

QY 81 PheGlyValAlaIleAspGlyIleAspGlyIleThrTyrAspAsnIleAlaPheGlnGlnAlaTyrAsn 100
 DB 291 TTGGAGCTTAAGGGTATGAGAAAACATATGATTAATATTGATTTGAGCAACATGAT 350
 QY 101 GluAlaCysSerSerArgThrProValGlnPheValAlaProValAsnIleAsnTyrIleu 120
 DB 351 GAAGCATGTTCAATCTAGAACACCTGTTCATTTGGTTCCTTAATAACAGAAATATCTT 410
 QY 121 LeuIleGlnIleIleThrPheSerGlyProCysArgSerSerIleSerValIleIlePheGly 140
 DB 411 CTCAGCAAAATCACCTTTTCAGTCCATGCAGATCTTCAATTCAGTAAAGATTTTTGA 470
 QY 141 SerIleuGlnAlaSerSerIleSerAspTyrIleAspAspArgGlyLeuTyrIleAlaPhe 160
 DB 471 TCCTTAGAGCATCTAGTAATAATTTCAAGCTCAAGATGAGAGGCTTTGAGATGCTTTT 530
 QY 161 AspSerValGlnAsnLeuValAlaGlyIleGlyIleThrIleAsnGlyAsnGlyVal 180
 DB 531 GATAGTGTCAAAATTTAGTTGTTGAGAGAGAGAACTATCATATGCAATGACAAAGTA 590
 QY 181 TrpTyrProSerSerCysValIleAsnIleSerIleuProCysArgAspAlaProThrAla 200
 DB 591 TGGTGGCAAGTTCTTGCAGAAATTAATAATCACTGCGCATGACAGAGATGACCAACGCGC 650
 QY 201 LeuThrPheTyrAsnCysIleAsnLeuIleValAsnIleuIleSerIleAsnAlaGln 220
 DB 651 TTAACTCTCGAATTGCAAAAATTTGAAAGTGAATATCTAAGAGTAAAGTAAATTCACAA 710
 QY 221 GlnIleHisIleIleAspPheGlnIleSerCysThrAsnValAlaAlaSerAsnLeuMetIleAsn 240
 DB 711 CAAATTCATCAAAATTTGAGTCAATGACCACTAAATGTTGATCTTCAAAATTTGATGATCAAT 770
 QY 241 AlaSerAlaIleSerProAsnThrAspGlyValAlaIleValSerAsnThrGlnTyrIleGln 260
 DB 771 GCTTCAGCAAAAGCCCAATATGATGAGTCAATGATCAATCAATCTCAATATTATTCAA 830
 QY 261 IleSerAspThrIleIleGlyIleThrGlyAspAspCysIleSerIleValIleSerGlySerGln 280
 DB 831 ATATCTGATCTATTAATTTGAGACAGATGATGATTAATTTCAATTTCTGATCTCA 890
 QY 281 AsnValGlnAlaThrAsnIleThrCysGlyProGlyHisGlyIleSerIleGlySerIleu 300
 DB 891 AATGTGAGGCAACCAATATTAATTTGATCTGATGATGATGATGATGATGATGATGATGAT 950
 QY 301 GlySerGlyAsnSerGlnAlaTyrValSerAsnValThrValAsnGlnAlaIleIle 320
 DB 951 GGAATCGGAATTCAGAACTTATGTCTTAATGTTACTGTAATGAGCCAAAATTTATC 1010
 QY 321 GlyAlaGlnAsnGlyValArgIleIleIleIleIleIleIleIleIleIleIleIleIleIle 340
 DB 1011 GGTGCCGAAATGAGATTTAGATCAAGACTTGTGCGAGGAGGATCTGACAGCTGAGCAAC 1070
 QY 341 IleIlePheLeuAsnValGlnIleGlnIleAspValIleIleIleIleIleIleIleIleIle 360
 DB 1071 ATCAAAATTTCTGAATGTGAAATGCAAGACGTTAGTATCCATATTAATAGCAACCAAAAC 1130
 QY 361 TyrCysAspArgValGlnProCysIleGlnGlnPheSerAlaValGlnValIleAsnVal 380
 DB 1131 TATTTGATCGATGTTGAACCATGATATACAAACAGTTTTCAGAGCTTCAAGTAAAGTGTG 1190
 QY 381 ValIleTyrGlnAsnIleIleGlyIleThrSerAlaThrIleValAlaIleIlePheAspCysSer 400
 DB 1191 GTGTATGAGAAATATCAAGGGCAAGTGCACAAAGATGGCCATTAATTTGATTTGACAGC 1250
 QY 401 ThrAsnPheProCysGlnGlyIleIleIleIleIleIleIleIleIleIleIleIleIleIle 420
 DB 1251 ACAAACTTTTCATGTGAGAGAAATTAATGAGAAATTAATTAATTAATTAATTAATTAATTAAT 1310
 QY 421 LysProSerGlnAlaThrCysIleAsnValHisPheAsnAsnIleGlnHisValIleThrPro 440
 DB 1311 AAACCATCAAGAGCTACGTGCAAAAATGTCATTTTAACAATGCTGAACAATGTTTACACCA 1370
 QY 441 HisCysThrSerLeuGlnIleSerGlnAspGlnAlaLeuIleuIleuIleuIleuIleuIleu 457

```
Db      1371 CACTGCACTTCACTAGAAATTTGAGAGATGAAGCTCTTTGTATTAATT 1421
RESULT 4
AA80487
ID      AA80487 standard; DNA; 1637 BP.
XX
XX      AA80487;
XX
XX      25-MAR-2003 (revised)
DT      20-NOV-1990 (first entry)
XX
XX      Plasmid clone pTom6.
DE
XX      Fruit ripening; polygalacturonase; pectin esterase; ss.
XX
XX      Synthetic.
OS
XX      Key      Location/Qualifiers
XX      CDS      51..1424
XX      FT      /tag= a
XX      FT      /label= polygalacturonase
XX
XX      EP271988-A.
XX
XX      22-JUN-1988.
XX
XX      06-NOV-1987; 87EP-00309853.
XX
XX      11-NOV-1986; 86GB-00026879.
XX
XX      (ICIL ) IMPERIAL CHEM IND PLC.
XX      (ZENB ) ZENBICA LTD.
XX
XX      Bridges IG, Schuch WW, Grierson D;
XX
XX      WPI; 1988-169271/25.
XX      P-FSDB; AAP80299.
XX
XX      Recombinant DNA comprising promoter and terminator sequences - useful in
XX      plants for altering ripening properties esp. in tomatoes.
XX
XX      Disclosure; Page ?; 22pp; English.
XX
XX      This cDNA clone, encoding polygalacturonase (PG), is used to produce
XX      antisense mRNA (with an inverted sequence to that of PG mRNA) which is
XX      inserted into a vector used to transform plants which thereafter have
XX      altered ripening properties. The inverted sequence and the PG mRNA form a
XX      double-stranded structure which inhibits expression of the PG mRNA. See
XX      also AA80488. (Updated on 25-MAR-2003 to correct PA field.)
XX
XX      Sequence 1637 BP; 598 A; 232 C; 290 G; 517 T; 0 U; 0 Other;
SQ
SQ
Alignment Scores:
pred. No.: 6.09e-219 Length: 1637
Score: 2390.00 Matches: 457
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0
US-10-691-374-2 (1-457) x AA80487 (1-1637)
QY      1 MetValIleGlnArgAsnSerIleLeuLeuIleIleIlePheAlaSerSerIleSer 20
Db      51 ATGGTATATCCAAAGGATAGTATTTCTCTCATATTAATTTTGGTTTCATCAATTTTCA 110
QY      21 ThrCysArgSerAsnValIleAspAspAsnLeuPheLeuGlnValIleAspAsnIleLeu 40
Db      111 ACTTGAGAGAGCAATGTTATGATGACAAATTTATTCAAACAAAGTTTATGATTAATTTCTT 170
QY      41 GlnGlnGlnPheAlaHisAspPheGlnAlaTyrLeuSerTyrLeuSerIleAsnIleGln 60
|||||
```

```
Db      171 GAACAAAGATTGGTCATGATTTTCAAGCTTATCTTCTTATTGACAAAAATATTGAA 230
QY      61 SerAsnAsnAsnIleAspIleValIleAspIleAsnGlnIleIleValIleAsnValIleuSer 80
Db      231 AGCAACAATTAATATTGACAAAGGTTGATATAAAATGGATTTAAAGTATTAATTAATCTTAGC 290
QY      81 PheGlnAlaIleGlnIleAspGlnIleAspGlnIleAspAsnIleAlaPheGlnGlnIleIlePheAsn 100
Db      291 TTGGAGCTTAAGGGTGAATGAGAAAAATATGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 350
QY      101 GlnAlaCysSerSerArgThrProValGlnPheValIleProIleAsnIleAsnIleuSerIleu 120
Db      351 GAAGCATGTTCAATCTAGAAACACTGTTCAATTTTGTTGTTCTTAATAAACAAGATTAATCTT 410
QY      121 LeuIleGlnIleThrPheSerGlnProCysArgSerSerIleSerValIlePheGlnIle 140
Db      411 CTCAGCAAAATCACTTTTCAGGTCACATGACATCTTCAATTTTCAGTAAAGATTTTTCGA 470
QY      141 SerIleGlnAlaSerSerIleIleSerAspTyrIleAspAspArgLeuTrpIleAlaPhe 160
Db      471 TCCTTAGAAGCATCTAGTAAATTTTCAGATCAACAAAGATAGAAAGGCTTTGGATGCTTTT 530
QY      161 AspSerValGlnAsnLeuValIleGlnIleGlnIleThrIleAsnGlnIleAsnGlnIleVal 180
Db      531 GATAGTGTTCAAAATTTAGTCTTTGAGAGAGAGAGACATCAATGCGCAATGCAACAAAGTA 590
QY      181 TrpTrpProSerSerCysIleValIleAsnIleSerIleuProCysArgAspAlaProThrAla 200
Db      591 TGTGGCCAAAGTTCTTGCAAAATTAATTAATCACTGCGCATGCGAGGATGACCAACGCGCC 650
QY      201 LeuThrPheThrAsnCysIleAsnLeuIleValIleAsnIleuSerIleAsnIleGln 220
Db      651 TTACCTCTTGAAATTCGAAAAATTTGAAAGGAAATTAATCTAAGAGTAAAGTAAATGACAA 710
QY      221 GlnIleHisIleIlePheGlnIleuSerCysThrAsnValIleAlaSerAsnIleuMetIleAsn 240
Db      711 CAATTCATATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 770
QY      241 AlaSerAlaIleSerProAsnThrAspGlnValIleAlaValSerAsnThrGlnIleGln 260
Db      771 GCTTCGCAAAAGGCCAAATCTGATGAGTCCATGATCAATCAATCAATCAATCAATCAATCAAT 830
QY      261 IleSerAspThrIleIleGlnIleThrGlnIleAspAspCysIleSerIleValSerGlnIle 280
Db      831 AATATCTGATATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 890
QY      281 AsnValGlnAlaThrAsnIleThrCysGlnProGlnIleGlnIleSerIleGlnIleuSerIleu 300
Db      891 AATGTCAGGCGCAAAATTTTCTTGCTCCAGGTCATGATTAATTAATTAATTAATTAATTAATTAAT 950
QY      301 GlnSerGlnIleAsnSerGlnAlaTyrValSerAsnValThrValAsnGlnAlaIleIleIle 320
Db      951 GGATCTGCAAAATTCAGAACTTATGATGCTTAATGTTACTGTAAATCAACCAAAATTTATTC 1010
QY      321 GlnAlaGlnIleAsnGlnValIleGlnIleThrGlnIleGlnIleGlnIleGlnIleAsnIle 340
Db      1011 GGATGCGGAAAAAGGATTAATGATTAATCAAGATTTGCGAGGAGATCTGGAACAAGCTACCAAC 1070
QY      341 IleIlePheLeuAsnValGlnIleMetGlnIleAspValIleTyrProIleIleIleAspGlnIle 360
Db      1071 ATCAATTTCTGAATGTTGGAATTCGAAGCGTTAAGTATCCATTAATTAATTAATTAATTAATTAAT 1130
QY      361 TyrCysAspAspValGlnIleProCysIleGlnIlePheSerAlaValGlnIleValIleAsnVal 380
Db      1131 TATTTGATCGAGTTGCAACCATGATATCAACAGTTTTCGAGCTTCAATGAAAAATATG 1190
QY      381 ValIleTyrGlnIleIleIleGlnIleThrSerIleAlaThrIleValIleIleIlePheAspCysSer 400
Db      1191 GGTATGAAATATCAAGGCGCAAGTGCACAAAGGAGGCGCAATTAATTAATTAATTAATTAATTAAT 1250
QY      401 ThrAsnPheProCysGlnIleIleIleMetGlnIleIleAsnIleValGlnIleuSerGlnIle 420
Db      1251 ACBAATTTCCATGTCAGAGATTAATTAATGAGAAATTAATTAATTAATTAATTAATTAATTAATTAAT 1310
```

QY 421 LysProSerGluAlaThrCysGlyAsnValHisPheAsnAsnAlaGluHisValThrPro 440
DB 1311 AACCATCAGAGGCTAGCTGCAAAATGCTCATTTTAACTGTAACATGTTTACACA 1370
QY 441 HisCysThrSerLeuGluLeuSerGluAspGluAlaLeuLeuYrtyrThr 457
DB 1371 CACTGCACCTTCACTAGAAATTTTCAGAGATGAAAGCTCTTTGTATATATAT 1421
RESULT 5
AAS20853
ID AAS20853 standard; DNA, 5822 BP.
XX AAS20853;
AC AAS20853;
XX 05-JUN-2002 (first entry)
DT
XX
DE Expression cassette, pFP-IRN1 useful for gene silencing.
XX
KW Inhibition of gene expression; expression cassette; inverted repeat; IR;
KW NOS; nopaline synthase; figwort mosaic virus promoter; FMV; hsp70; PG;
KW plant heat shock 70; tomato; polygalacturonase; disease resistance;
KW flavour; nutritional characteristic; plant; gene silencing; pFP-IRN1;
KW regulating gene expression; mutant; ds.
XX
OS Lycopersicon esculentum.
OS Agrobacterium tumefaciens.
OS Figwort mosaic virus.
OS Synthetic.
OS Chimeric.
OS
PN WO200214472-A2.
XX
PD 21-FEB-2002.
XX
PF 14-AUG-2001; 2001WO-US025538.
XX
PR 15-AUG-2000; 2000US-0225508P.
PR 07-AUG-2001; 2001US-00924197.
XX
XX
PA (DNAP) DNA PLANT TECHNOLOGY CORP.
PI Guttersen N, Oeller P;
XX
XX WPI; 2002-257599/30.
DR
XX
PT Reducing the expression of a target gene in a cell, comprises expressing
PT in the cell an expression cassette comprising a promoter operably linked
PT to a sense or antisense targeting sequence and an inverted repeat of a
PT nopaline synthase gene.
XX
PS Claim 51; Page 37-39; 39pp; English.
XX
XX The present invention relates to an improved method for inhibiting the
XX expression of a target gene in a cell, by expressing in the cell an
XX expression cassette comprising a promoter operably linked to a sense or
XX antisense targeting sequence having substantial identity to a subsequence
XX of the target gene, and an inverted repeat (IR) of a subsequence of a NOS
XX (nopaline synthase) gene, where the IR is unrelated to the targeting
XX sequence; The expression cassette, pFP-IRN1 is constructed using a
XX figwort mosaic virus (FMV) promoter in which the 5'-untranslated leader
XX (5'-UTL) is derived from a plant heat shock 70 (hsp70) gene, the tomato
XX polygalacturonase (PG) gene and an IR of the terminator of the
XX Agrobacterium tumefaciens NOS gene. The method is useful for regulating
XX expression of endogenous genes and transgenes, e.g. to regulate
XX flavour, protein or nutritional characteristics. The improved gene
XX silencing construct is used in functional genomics to determine the
XX effect of regulating gene expression of a selected endogenous gene or
XX transgene. The method is simple and rapid, and is suitable for high-
XX throughput studies. Multiple transgenic constructs all containing the
XX same repeat element can be silenced at the same time, since the initial
XX silencing trigger mediated through the inverted repeat region will apply

CC to all of the transcripts. The present sequence represents the expression
CC cassette, pFP-IRN1
XX
SQ Sequence 5822 BP; 1689 A; 1275 C; 1292 G; 1566 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1 49e-163 Length: 5822
Score: 1820.00 Matches: 345
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.71% Mismatches: 0
Query Match: 76.15% Indels: 0
DB: 6 Gaps: 0
US-10-691-374-2 (1-457) x AAS20853 (1-5822)
QY 112 ValValProLysAsnLysAsnTyrLeuLeuLysGluLeuThrPheSerGlyProCysArg 131
DB 1387 ATGGTTCCTMAAAACAGATATATCTTCTCAAGCAATCACTTTTCAAGTCATGCA 1446
QY 132 SerSerLeSerValLysLePheGlySerLeuGluAlaSerSerLysLeSerAspTyr 151
DB 1447 TCTTCTATTTCACTAAAGATTTTGGATCTTGAAGCATCTAGTAAATTTTCAGACTAC 1506
QY 152 LysAspArgArgLeuTyrPheAlaPheAspSerValGlnSerLeuValValGlyGly 171
DB 1507 AAAGATAGAGAGCTTTGGATTTGATGTTGATGTTCAAAATTTAGTTTGGAGAGGA 1566
QY 172 GlyThrLeaAsnGlyAsnGlyGlnValTyrThrProSerSerCysLysLysLeuLysSer 191
DB 1567 GGAATCTATCAATGGCAATGACCAAGATGGGCGCAAGTTCTTGCAAAATTAATCA 1626
QY 192 LeuProCysArgAspAlaProThrAlaLeuThrPheThrAsnCysLysAsnLeuLysVal 211
DB 1627 CTGCCATGACAGGATGACCAACAGCCTTAACCTTGGATTTGCAAAATTTGAAGTG 1686
QY 212 AsnAsnLeuLysSerLysAsnAlaGlnGlnHisLeLysPheGluSerCysThrAsn 231
DB 1687 AATATCTTAAGAGTATTAAGATGCAACCAATTCATATCAATTTAGTCATGCACTAAT 1746
QY 232 ValValAsSerAsnLeuMetLeaAsnAlaSerAlaLysSerProAsnThrAspGlyVal 251
DB 1747 GTTGTACCTTCAATTTGATGATCAATGCTTACGCAAAAGCCCAATCTGATGAGTC 1806
QY 252 HisValSerAsnThrGlnTyrTyrLeuGlnLeSerAspThrLeLeGlyThrGlyAsp 271
DB 1807 CATGTATCAATATCTCAATATATTCATATCTGATCATATTTGGAACAGGTATGAT 1866
QY 272 CysLeSerLeValSerGlySerGlnAsnValGlnAlaThrAsnLeThrCysGlyPro 291
DB 1867 TGTATTTCAATTTGTTCTGATCTCAAAATGCGAGCCCAAAATATTAATCTGTGCTCA 1926
QY 292 GlyHisGlyLeSerLeLeGlySerLeuGlySerGlyAsnSerGluAlaTyrValSerAsn 311
DB 1927 GGTCAAGGATTAAGATTTGGAAGCTTGAATCTGGAATTCGAAGCTTATGTGCTAAT 1966
QY 312 ValThrValAsnGluAlaLysLeLeLeGlyAlaGluAsnGlyValArgGlyLeuThrTyr 331
DB 1987 GTTACTGTAAATGAAGCCAAATTTATCGTCCCAAAATGAGATTAGATCAAAACTTGG 2046
QY 332 GlnGlyGlySerGlyGlnAlaSerAsnLeLysPheLeuAsnValGluMetGlnAspVal 351
DB 2047 CAGGAGGATTCGACCAAGCTAGCAACATCAATTTCTGATGTGGAATGCAAGAGCTT 2106
QY 352 LysTyrProLeuLeLeLeAspGluAsnTyrCysAspArgValGluProCysLeGlnGln 371
DB 2107 AAGTATCCCAATATTAATTAAGACCAAACTATTTGATGATGAATGAACCATATTAACAAG 2166
QY 372 PheSerAlaValGlnValLysAsnValTyrGluAsnLeLysGlyThrSerAlaThr 391
DB 2167 TTTTCAGAGTTCAAGGAAAAATGTGTATGAGATATCAAGGCAATGTCACCA 2226
QY 392 LysValAlaLeuLysPheAspCysSerThrAsnThrProCysGluGlyLeuLeuMetGlu 411

Db	2227	AAGGTGGCCATTAATTTGATTCGACGACCAACTTTCATGTGAAGAAATTATATGAG	2286		
Oy	412	AenIleuLeuValGlyGluSerGlyPyrProSerGluAlaThrCysIleAsnValHis	431		
Db	2287	AATATTAATTTAGTAGGGGAAAGTGGAACCATTCAGGGCTAGCTGCAAAAATGTCCAT	2346		
Oy	432	PhenAlaGluValThrProHisCysThrSerLeuGluIleSerGluAspGlu	451		
Db	2347	TTTAAACAATGCTGGAACATGTATACACCACTGCACCTTCACTAGAAATTTTCAGAGATGAA	2406		
Oy	452	AlaIleuLeuTyrAsnTyr	457		
Db	2407	GCTCTTTTGTATTAATTTAT	2424		
RESULT 6					
ADG87885					
ID	ADG87885	standard; cDNA; 1296 BP.			
XX	ADG87885;				
DT	22-APR-2004	(first entry)			
XX					
De	A. thaliana RPP7/RPP8-upregulated pathogen infection-related gene #327.				
XX					
KW	Pathogen infection-related gene; plant; Peronospora parasitica; defence mechanism; RPP7; RPP8; pathogen resistance; transgenic plant; oomycete; fungus; bacterium; virus; nematode; insect; aphid; gene; ss.				
XX					
OS	Arabidopsis thaliana.				
XX					
PN	WO200222675-A2.				
PD	21-MAR-2002.				
XX					
PF	14-SEP-2001; 2001WO-US028506.				
XX					
PR	15-SEP-2000; 2000US-0232778P.				
PR	22-JUN-2001; 2001US-0300183P.				
XX					
PA	(SYGN) SYNGENTA PARTICIPATIONS AG.				
PA	(UTNC-) UNIT NORTH CAROLINA.				
PA	(GLAZ/) GLAZEBROOK J.				
PA	(WANG/) WANG X.				
PA	(DANG/) DANGJ J L.				
PA	(EULG/) EULGEM T.				
PA	(ZHUT/) ZHU T.				
XX					
PI	Glazebrook J, Wang X, DangJ JL, Eulgem T, Zhu T;				
XX					
XX					
XX					
PT	Novel isolated polynucleotide, useful for conveying pathogen resistance to plants, and for identifying plants infected with a pathogen.				
XX					
PS	Claim 3; SEQ ID NO 327; 605bp; English.				
XX					
CC	The invention relates to 691 Arabidopsis thaliana genes (ADG87559--ADG87557) whose expression is altered in response to pathogen infection, or to homologues of these genes from other plants or fungi, especially from maize, soybean, barley, alfalfa, sunflower, canola (oilseed rape), cotton, peanut, sorghum, tobacco, sugarbeet, rice or wheat. The expression of genes of the invention was upregulated or downregulated in Arabidopsis plants infected with the oomycete Peronospora parasitica, indicating that they play a role in defence mechanisms. The genes of the invention are regulated by RPP7 or RPP8 which act via unconventional signalling cascades, or by the RPP4-dependent pathway. The invention also relates to polypeptides encoded by the pathogen infection-related genes; promoter motifs from pathogen infection-related genes (ADG88243-ADG88227); expression cassettes, host cells and pathogen-resistant transgenic plants and their progeny comprising a polynucleotide of the invention; and a method of identifying a plant cell infected with a pathogen. The polynucleotide sequences and methods of the invention are useful for identifying plants infected with a pathogen, and for conferring				

resistance to pathogens such as oomycetes, fungi, bacteria, viruses, nematodes and insects (e.g., aphids). The present sequence represents an Arabidopsis thaliana gene whose expression is altered in response to Peronospora parasitica infection. Note: The sequence data for this patent did not form part of the published specification, but was obtained in electronic format directly from WIPO at http://wipo.int/pub/published_pct_sequences.

Sequence 1296 BP; 424 A; 244 C; 293 G; 335 T; 0 U; 0 Other;

Alignment Scores:

Magnum 2000:	
Pred. No.:	4.98e-98
Score:	1127.00
Length:	1296
Matches:	223

Score:	112.700	Matches:	225
Percent Similarity:	65.33%	Conservative:	71
Best Local Similarity:	49.56%	Mismatches:	116

Best Local Similarity:	45.58	MissMatches:	110
Query Match:	47.15	Indels:	40
NR:	6	Gaps:	6

MS-10-691-374-2 (1-457) x ADG87885 (1-1296)

7 serr[e]l[e]n[e]l[e]t[e]t[e]phea

Dh
75 GCGGCTTTCCTTAACGCTTCTGGTGAGTAAAGCTTTGAGTAGCAACGTTT 84

|||||
:::| | :::
|||

[illegible]

144

1. **Introduction**
 2. **Background**
 3. **Methodology**
 4. **Results**
 5. **Discussion**
 6. **Conclusion**
 7. **References**
 8. **Appendix**
 9. **Index**
 10. **Table of Contents**
 11. **Figure 1**
 12. **Figure 2**
 13. **Figure 3**
 14. **Figure 4**
 15. **Figure 5**
 16. **Figure 6**
 17. **Figure 7**
 18. **Figure 8**
 19. **Figure 9**
 20. **Figure 10**
 21. **Figure 11**
 22. **Figure 12**
 23. **Figure 13**
 24. **Figure 14**
 25. **Figure 15**
 26. **Figure 16**
 27. **Figure 17**
 28. **Figure 18**
 29. **Figure 19**
 30. **Figure 20**
 31. **Figure 21**
 32. **Figure 22**
 33. **Figure 23**
 34. **Figure 24**
 35. **Figure 25**
 36. **Figure 26**
 37. **Figure 27**
 38. **Figure 28**
 39. **Figure 29**
 40. **Figure 30**
 41. **Figure 31**
 42. **Figure 32**
 43. **Figure 33**
 44. **Figure 34**
 45. **Figure 35**
 46. **Figure 36**
 47. **Figure 37**
 48. **Figure 38**
 49. **Figure 39**
 50. **Figure 40**
 51. **Figure 41**
 52. **Figure 42**
 53. **Figure 43**
 54. **Figure 44**
 55. **Figure 45**
 56. **Figure 46**
 57. **Figure 47**
 58. **Figure 48**
 59. **Figure 49**
 60. **Figure 50**
 61. **Figure 51**
 62. **Figure 52**
 63. **Figure 53**
 64. **Figure 54**
 65. **Figure 55**
 66. **Figure 56**
 67. **Figure 57**
 68. **Figure 58**
 69. **Figure 59**
 70. **Figure 60**
 71. **Figure 61**
 72. **Figure 62**
 73. **Figure 63**
 74. **Figure 64**
 75. **Figure 65**
 76. **Figure 66**
 77. **Figure 67**
 78. **Figure 68**
 79. **Figure 69**
 80. **Figure 70**
 81. **Figure 71**
 82. **Figure 72**
 83. **Figure 73**
 84. **Figure 74**
 85. **Figure 75**
 86. **Figure 76**
 87. **Figure 77**
 88. **Figure 78**
 89. **Figure 79**
 90. **Figure 80**
 91. **Figure 81**
 92. **Figure 82**
 93. **Figure 83**
 94. **Figure 84**
 95. **Figure 85**
 96. **Figure 86**
 97. **Figure 87**
 98. **Figure 88**
 99. **Figure 89**
 100. **Figure 90**
 101. **Figure 91**
 102. **Figure 92**
 103. **Figure 93**
 104. **Figure 94**
 105. **Figure 95**
 106. **Figure 96**
 107. **Figure 97**
 108. **Figure 98**
 109. **Figure 99**
 110. **Figure 100**
 111. **Figure 101**
 112. **Figure 102**
 113. **Figure 103**
 114. **Figure 104**
 115. **Figure 105**
 116. **Figure 106**
 117. **Figure 107**
 118. **Figure 108**
 119. **Figure 109**
 120. **Figure 110**
 121. **Figure 111**
 122. **Figure 112**
 123. **Figure 113**
 124. **Figure 114**
 125. **Figure 115**
 126. **Figure 116**
 127. **Figure 117**
 128. **Figure 118**
 129. **Figure 119**
 130. **Figure 120**
 131. **Figure 121**
 132. **Figure 122**
 133. **Figure 123**
 134. **Figure 124**
 135. **Figure 125**
 136. **Figure 126**
 137. **Figure 127**
 138. **Figure 128**
 139. **Figure 129**
 140. **Figure 130**
 141. **Figure 131**
 142. **Figure 132**
 143. **Figure 133**
 144. **Figure 134**
 145. **Figure 135**
 146. **Figure 136**
 147. **Figure 137**
 148. **Figure 138**
 149. **Figure 139**
 150. **Figure 140**
 151. **Figure 141**
 152. **Figure 142**
 153. **Figure 143**
 154. **Figure 144**
 155. **Figure 145**
 156. **Figure 146**
 157. **Figure 147**
 158. **Figure 148**
 159. **Figure 149**
 160. **Figure 150**
 161. **Figure 151**
 162. **Figure 152**
 163. **Figure 153**
 164. **Figure 154**
 165. **Figure 155**
 166. **Figure 156**
 167. **Figure 157**
 168. **Figure 158**
 169. **Figure 159**
 170. **Figure 160**
 171. **Figure 161**
 172. **Figure 162**
 173. **Figure 163**
 174. **Figure 164**
 175. **Figure 165**
 176. **Figure 166**
 177. **Figure 167**
 178. **Figure 168**
 179. **Figure 169**
 180. **Figure 170**
 181. **Figure 171**
 182. **Figure 172**
 183. **Figure 173**
 184. **Figure 174**
 185. **Figure 175**
 186. **Figure 176**
 187. **Figure 177**
 188. **Figure 178**
 189. **Figure 179**
 190. **Figure 180**
 191. **Figure 181**
 192. **Figure 182**
 193. **Figure 183**
 194. **Figure 184**
 195. **Figure 185**
 196. **Figure 186**
 197. **Figure 187**
 198. **Figure 188**
 199. **Figure 189**
 200. **Figure 190**
 201. **Figure 191**
 202. **Figure 192**
 203. **Figure 193**
 204. **Figure 194**
 205. **Figure 195**
 206. **Figure 196**
 207. **Figure 197**
 208. **Figure 198**
 209. **Figure 199**
 210. **Figure 200**
 211. **Figure 201**
 212. **Figure 202**
 213. **Figure 203**
 214. **Figure 204**
 215. **Figure 205**
 216. **Figure 206**
 217. **Figure 207**
 218

[illegible]

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

[illegible][illegible][illegible][illegible]

20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525
 526
 527
 528
 529
 530
 531
 532
 533
 534
 535
 536
 537
 538
 539
 540
 541

DB 262 AAGAAAGCAI GGAAGAAGCAIG I CAAACAAA I GGAG I AAC AAC I I C I I CGA I CC C AAA 32

116 AsnlybAsnlytLeuLeuLybGlnIleInrPheserGlyProCybArGserSerIleSer 153

D5 322 GGAAGACTTATCTCCTAAGTCTACCTCGATTAGAGGCCAATGCAATCTTACGTAAC 381

QY 136 ValValIlePheGlySerLeuGluAlaSerSerLysIleSerAspTyrLysAspArg 155

Db 382 TTTCAGATCCTAGGCACTTTATCAGCATCTACGAAACGTTCAGATTACAAAGACAAAAC 441

156 LeuTrpIleAlaPheaspSerValGlnasnLeuValValGlyGlyGly-----GlyThr 173

Db 442 CATTGGCTTATCTAGAGGACGTTACAATCTATCAATCGACGGTGGCTCGACGGGAATT 501

QY 174 IleAsnGlyAsnGlyGlnValIrrTPProSerSerCysIleAsnLysSerLeuPro 193

Db 502 ATTAATGCCAACGAAACCTGCTGGCAGAACTCATGCAAAATCGACAAATCTAAGCCA 561

194 CysArgAspAlaProThrAlaLeuThrPheTyrAsnCysValAsnLeuValAsnVal 213

Db 562 TGCACAAAGCTCCACGGCTTACTTTATACAATTTAAGAAATTGAATGTGAAGAT 621

214 LeuLysSerLysAsnAlaGlnGlnIleHisIleLeuPheGluSerCysThrAsnValVal 233

Db 622 CTGAGGGTGAATAATGCCGACGACAGTTCAGATTCAATTGAGAATGCACAAAGTTGAA 681

234 AlaSerArgLeuMetIleAsnAlaSerAlaIleValSerProArgThrAspGlyValHisVal 253

Dh 682 GTTACTAATGTTGAGATCACTGCTCCGGCGATAGTCCCAACACAGATGATCCATATC 741

354 servant[et]le[et]v[et]le 273


```
QY 234 A1aseAsnleuMetIleAsnAlaseRAlaIySeSerProAsnthrAspGlyValHisVal 253
Db 682 GTTAGTAATGTTAGATCACTGCTCCGGGCGATATGTCACAAACAAATGATATCCATATC 741
QY 254 SerAsnThrGlnTyrlleGlnIleSerAspThrIleGlyThrGlyAspAspCysIle 273
Db 742 ACTAATATCTCAAAACATTGAGCTCCAACTCAATATGGAACAGGTGATGATTTTAA 801
QY 274 SerIleValSerGlySerGlnAsnValGlnAlaThrAsnIleThrCysGlyProGlyHis 293
Db 802 TCCATTAGAGATGGAACGCAAAATCTTCAATCTTCAATCTTCAATCTTCAATCTTCAAT 861
QY 294 GlyIleSerIleGlySerIleGlySerGlyAsnSerGlnAlaThrValSerAsnValThr 313
Db 862 GGGATGAGCTTTGGAGCTTTGGAGCTTTGGAGCTTTGGAGCTTTGGAGCTTTGGAGCTTT 921
QY 314 ValAsnGlnAlaIleValIleIleGlyAlaGlnAsnGlyValArgIleIleThrTrpGlnGly 333
Db 922 GTGGATGAGCTTTAGTCTCTGAGATGACAAATGAGATTAGATTAGATTAGATTAGATTAG 981
QY 334 GlySerGlyGlnAlaSerAsnIleIleIleIleIleIleIleIleIleIleIleIleIle 353
Db 982 GGATTCGAACTGCGCAAGAACTTAAATTTCAAAATTTCAAAATTTCAAAATTTCAAAAT 1041
QY 354 ProIleIleIleAspGlnAsnTyIleCysAspArgValGluProCysIleGlnIlePheSer 373
Db 1042 CCGATATATATCAACCAAGGACTATGCGCAAG--GACAAATGCAAAATCCAAAGCTG 1098
QY 374 A1aValGlnValIleAsnValValIleGlnAsnIleIleGlyThrSerAlaThrIleVal 393
Db 1099 GCAGTGCAGTGAAGAAACGTTGTCACAAATATCTGTCGATGCGAGCTGACGATGTCG 1158
QY 394 A1aIleIleIleAspCysSerThrAsnIleIleIleIleIleIleIleIleIleIleIle 413
Db 1159 GCGATTACCTGTCGATGCGAGGAAATTCATGTCGATGTCGATGTCGATGTCGATGTCG 1218
QY 414 AsnIleValGlyIleGlySerGlyIleProSerGlnAlaThrCysIleAsnValHisPheAsn 433
Db 1219 AAAATTAAGGA-----GGAAACGCTTCTTGCAAAATGCCAATGTTAA 1263
QY 434 AsnAlaGlnHisValThrProHisCysThr 443
Db 1264 NATCAAGCACCGCTTCTCTCAATGCTCT 1293
RESULT 8
AAC46786 standard; DNA; 1393 BP.
ID AAC46786;
AC AAC46786;
XX
XX
XX 18-OCT-2000 (first entry)
DT
XX
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 51398.
DE
XX
XX Hybridization assay; genetic mapping; gene expression control;
KM protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
XX Arabidopsis thaliana.
OS
XX
XX PN EP1033405-A2.
PD
XX 06-SEP-2000.
XX
XX PF 25-FEB-2000; 2000EP-00301439.
XX
XX PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0128845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-013256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
```

PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160777P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	28-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161922P.
PR	28-OCT-1999;	99US-0161933P.
PR	29-OCT-1999;	99US-0162142P.

Alignment Scores:	
Pred. No.:	5 49e-98
Score:	1127.00
Percent Similarity:	65.33%
Best Local Similarity:	49.56%
Query Match:	47.15%
DB:	3
US-10-691-374-2 (1-457)	x AAC46786 (1-1393)
	Length: 1393
	Matches: 223
	Conservative: 71
	Mismatches: 16
	Indels: 40
	Gaps: 6

[illegible]

```

QY 174 11eaagc1yasn1ygl1nval1tpttpproserserCyulys11easnyserleuPro 193
DB 599 ATTAAATGCGAAACCGAAACCTGGTGCGAACTCGAAATCGAAATCTTAAGCCA 658
QY 194 CyArGAspAlarProthralaleuthrPheTPAspCyelysAsnleu1yValaAsn 213
DB 659 TGCACAAAAGCTCCAAACGGCTCTTACTTATATACAAATTTGAAGATGGAAGAA 718
QY 214 Leu1ySer1yAsn1a1a1n1n1n1n1n1n1n1n1n1n1n1n1n1n1n1n1n1n1 233
DB 719 CTGAGGGTGAATAATCGCGACAGATTCAATTCAATGAGAAATGCAACAAAGTTGA 778
QY 234 AlAsenAsnleu1n1e1eas1n1a1y1s1erProAsn1n1n1n1n1n1n1n1n1 253
DB 779 GTTAAATGTTGAAGATCACTCTCGCGGAGATGTCACAAACAAATGTAATTCAT 838
QY 254 SerAsn1n1n1n1n1n1n1n1n1n1n1n1n1n1n1n1n1n1n1n1n1n1n1 273
DB 839 ACTAATATCTAAACATTCAGATCTCACTCAATATGCAATGCAAGATGATGTA 898
QY 274 Ser11e1a1s1er1y1s1er1n1a1n1a1n1a1n1a1n1a1n1a1n1a1n1a 293
DB 899 TCATATGAGATGAGACGCAAAATCTTCAAAATCTTGAATTAATTCGCGCCGCT 958
QY 294 Gly11e1s1er11e1y1s1er1e1n1y1s1er1y1a1n1s1er1y1a1n1s1 313
DB 959 GGAATCAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 1018
QY 314 Vala1n1a1n1a1y1a1y1e1e1y1a1n1a1n1a1n1a1n1a1n1a1n1a1n 333
DB 1019 GTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1078
QY 334 Gly1s1er1y1a1n1a1s1er1n1e1y1s1er1y1a1n1a1n1a1n1a1n1a 353
DB 1079 GATTCAGAACTGCCAAGAACTTAATTTCAAAATATTCGATGAGAAACGTCAGA 1138
QY 354 Pro11e11e1e1a1n1n1n1n1n1n1n1n1n1n1n1n1n1n1n1n1n1n1n1 373
DB 1139 CCATATATATCAACCAAGATGATGATGATGATGATGATGATGATGATGATGAT 1195
QY 374 AlaVala1n1a1y1a1n1a1n1a1n1a1n1a1n1a1n1a1n1a1n1a1n1a 393
DB 1196 GCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1255
QY 394 Ala11e1y1a1n1a1n1a1n1a1n1a1n1a1n1a1n1a1n1a1n1a1n1a 413
DB 1256 GCGATATACGTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 1315
QY 414 Asn1e1u1a1n1a1n1a1n1a1n1a1n1a1n1a1n1a1n1a1n1a1n1a 433
DB 1316 AAAATAAAGGA-----GGAAACAGCTTCTTGCAAAATGCCATGTTAA 1360
QY 434 Asn1a1a1n1a1n1a1n1a1n1a1n1a1n1a1n1a1n1a1n1a1n1a1n 443
DB 1361 AATCAAGCAACCGTTCTCTTAATGCTCT 1390

```

```

PD 03-JAN-2003.
XX
XX 22-JUN-2001; 2001WO-1B001105.
XX
XX 22-JUN-2001; 2001WO-1B001105.
XX
XX (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y,
XX Katagiri F, Qian Y, Tao Y, Whitcham S, Xie Z, Zhu T, Zou G,
XX WPI, 2003-175290/17.
XX
PT Identifying at least one gene involved in plant resistance or response to
PT pathogenic infection for conferring resistance or tolerance to a plant to
PT bacterial, fungal or viral infection by determining or detecting plant
PT gene expression.
XX
PS Claim 6; SEQ ID NO 554; 899pp; English.
XX
XX The present invention relates to a method (M1) for identifying genes
XX involved in plant resistance or response to pathogenic infection. M1
XX comprises identifying a gene whose expression is significantly altered in
XX the incompatible interaction of plant gene expression relative to
XX expression of the gene in an uninfected plant, in a mutant plant that
XX does not express a gene associated with response to pathogenic infection,
XX or in a corresponding incompatible or compatible interaction. (M1) is
XX useful for conferring resistance to resistance or tolerance to a plant to
XX bacterial, fungal or viral infection. The present sequence was used to
XX illustrate the invention.
XX
SQ Sequence 1280 BP; 401 A; 257 C; 292 G; 330 T; 0 U; 0 Other;
XX
Alignment Scores:
Prid. No.: 6,19e-97 Length: 1280
Score: 115.50 Matches: 222
Percent Similarity: 66.37% Conservative: 74
Beet Local Similarity: 49.78% Mismatches: 117
Query Match: 46.67% Indels: 33
DB: 8 Gaps: 7
XX
US-10-691-374-2 (1-457) x ADA68538 (1-1280)
QY 6 Asn1e1u1a1n1a1n1a1n1a1n1a1n1a1n1a1n1a1n1a1n1a1n1a1n 25
DB 16 AACCTGTAAACCGTTTCCATATATGAGCTCTTGTGATGTCATGATGAAGCTTA 75
QY 26 Val11eAspAsnleuPhe1y1a1n1a1n1a1n1a1n1a1n1a1n1a1n1a1n1a 40
DB 76 AGAATCAGCCCTAAATGATATATACATTTTAAAGTTCAATCCATAGCTTAATC 135
QY 41 Glu1n1a1n1a1n1a1n1a1n1a1n1a1n1a1n1a1n1a1n1a1n1a1n1a1n 60
DB 136 AAGCAAGA-----GAGACATCACCGGATGAGAAAGTTTGAAGACCTCTTGCG 189
QY 61 SerAsnAsnAsn11eAsp1y1a1n1a1n1a1n1a1n1a1n1a1n1a1n1a1n 80
DB 190 ACTCCAAAC-----ACCGTAGTGTCTGAT 216
QY 81 PheGly1a1a1y1a1n1a1n1a1n1a1n1a1n1a1n1a1n1a1n1a1n1a1n1a 100
DB 217 TTGAGACTAAAGAGATGAGAAACCGATGACAGCAGCGCTTCGTGAATCGTGAAG 276
QY 101 Glu1a1Cy1s1er1a1n1a1n1a1n1a1n1a1n1a1n1a1n1a1n1a1n1a1n 120
DB 277 AAAGCATGTTCTTCAATGAGATGATGATGATGATGATGATGATGATGATGATG 336
QY 121 Leu1y1Gln11e1n1a1n1a1n1a1n1a1n1a1n1a1n1a1n1a1n1a1n1a 140
DB 337 CTTAAGCTTAATCAATTAATGATGATGATGATGATGATGATGATGATGATGAT 396
QY 141 Ser1e1u1a1s1er1y11e1s1er1y11e1s1er1y11e1s1er1y11e1s1er 160

```

```

Db 397 ACCGTATCCGCATCTCAAAAACGATCGATTACAAAGATATCGCAATATGTAATGTTT 456
Qy 161 AAGServalGlnAsnLeuValGlyGly-----GlyThrIleAsnGlyAsnGly 178
Db 457 GATGGGCTTAAACATCTATCATCGATGAGGCGACACTGGGCTTGAGACGAAACCGC 516
Qy 179 GlnValThrProSerCysGlyIleAsnLysSerLeuProCysArgAspAlaPro 198
Db 517 GAACCTGGTGGCAAAACTCATGCAACGAAACAAAGCT-----555
Qy 199 ThrAlaLeuThrPheThrAsnCysIlyAsnLeuLysValAsnAsnLeuLysSerLysAsn 218
Db 556 AAGGCTCTTACTTTTACAACTCGAAAGCTGATGAGAGAACTAGAGTGAAGAAAT 615
Qy 219 AlaGlnGlnIleHileleypheGlySerCysThrAsnValAlaSerAsnLeuMet 238
Db 616 GCACAGCAGATCAGATTTCGATTAAAAATGCTCAACGATCAGGTCTCTAATGCTGTG 675
Qy 239 IleaAsnAlaSerAlaLysSerProAsnThrAspGlyValHileValSerAsnThrGlnTyr 258
Db 676 GTAACGCGCGCTCGGATATGCTTAAACCGATGTTATCATATCATCAACCCCAAAAC 735
Qy 259 IlegGlnIleSerAspThrIleIlegIlyThrGlyAspAspCysIleSerIleValSerGly 278
Db 736 ATTGAGTCTCGAATCATATGGAACAGCGCATGATGATGATATCATATGAAAGTGA 795
Qy 279 SerGlnAsnValGlnAlaThrAsnIleThrCysGlyProGlyValIleGlyIleSerIleGly 298
Db 796 TCACAAAATGTTCAAAATGATATGATATGCTGCGGTCCGCTCAGGTATCAGATTTGGG 855
Qy 299 SerLeuGlySerGlyAsnSerGlnAlaLysValSerAsnValThrValAsnGlnAlaLys 318
Db 856 AGCTTGGAGATGACAATTCAAAGCTTTTGTCTCAGGCGTGACTGTGATGCTGCTAAG 915
Qy 319 IleIlegIlyAlaGlnAsnGlyValArgIleIleThrIleGlnGlySerGlyGlnAla 338
Db 916 CTTTCGGTACAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 975
Qy 339 SerAsnIleLysPheLeuAsnValGlnMetGlnAspValIleLysTyrProIleIleLeuAsp 358
Db 976 AGCAATATATATTTCAAAACATTCAGATGATATGATGATGATGATGATGATGATGATGAT 1035
Qy 359 GlnAsnIlyCysAspAspValGlnProCysIleGlnGlnPheSerAlaValGlnValLys 378
Db 1036 CAGACTACTGCGCAAG--AGCAATGCACTACAGAGAAATCCGCGTCCAAAGTGAAG 1092
Qy 379 AsnValValIlyGlnAsnIleLysGlyThrSerAlaThrIlyValAlaIleLysPheAsp 398
Db 1093 AAGCTGTTATCCGGGACATPAGTGGACATGAGGCATCGGAAACGCAATTAACGTTTAAAC 1152
Qy 399 CysSerThrAsnAspPheProCysGlyGlyIleIleMetGlnAsnIleAsnLeuValGlyGln 418
Db 1153 TGGAGCAAGAACTATCATGCAAGAAATTTGCTTGACAGAGTGAACCTTAAAGAA--- 1209
Qy 419 SerGlyLysProSerGlnAlaThrCysLysAsnValHilePheAsnAsnAlaIleuHileVal 438
Db 1210 -----GGAAGCAACTTCCACCAATGCTTAATGCTGTTAAGAGACGCTGT 1257
Qy 439 ThrProHisCysThrSer 444
Db 1258 CTGCTCAGTGCMACTCC 1275

```

```

XX Dehiscence zone; endo-polygalacturonase; seed shattering; oilseed rape;
XX transgenic plant; antisense; ribozyme; ds.
XX Brassica napus; cv. Topaz.
FH Key Location/Qualifiers
FT CDS /tag= a
FT /tag= b
FT sig_peptide /tag= b
FT mat_peptide /tag= b
FT /tag= c
FT primer_bind complement(821..837)
FT /tag= d
FT /note= "primer PG3"
FT complement(884..900)
FT primer_bind /tag= e
FT /note= "primer PG3"
FT /note= "primer PG3"
FT /tag= f
FT primer_bind /note= "primer PG2"
FT /tag= g
FT primer_bind /tag= g
FT /note= "primer PG5"
FT /tag= h
FT /note= "base 1439 is given as n in the specification"
PN WO9713865-A1.
PD 17-APR-1997.
PF 04-OCT-1996; 96WO-EP004313.
XX 06-OCT-1995; 95EP-00402241.
PR 08-DEC-1995; 95EP-00203328.
XX (PLBZ ) PLANT GENETIC SYSTEMS NV.
PI Ulyakov P, Child R, Van Onckelen H, Prinsen E, Borchardt B;
PI Sander L, Petersen M, Bundgaard Poulsen G, Botterman J;
DR WP1, 1997-235901/21.
XX Transgenic plant containing dehiscence zone selective chimeric gene - has
PT modified dehiscence properties, especially delayed pod dehiscence.
FT Claim 6; Page 48; 65pp; English.
PS
XX cDNA clone X (AAT63603) includes a coding sequence for oilseed rape
CC dehiscence zone (DZ)-selective endopolygalacturonase (PG). DZ cDNA was
CC subjected to PCR amplification using primers (see also AAT63605-08) based
CC on conserved regions of PG amino acid sequences. PG-related clones (see
CC also AAT63609-14) were identified, of which only PG35-8 was specific to
CC the DZ. This clone was used to screen a DZ-selective cDNA library prepd.
CC from poly-A+ RNA isolated from the DZ 6 wk after anthesis, yielding clone
CC X. PG35-8 was also used to screen a genomic library to identify the DZ-
CC selective PG gene promoter region (AAT63604). These sequences can be
CC utilised in novel chimeric genes to modify the dehiscence properties of
CC transgenic plants, partic. the pod dehiscence properties of B. napus.
CC This can improve the seed yield from the transformed plants. (Updated on
CC 25-MAR-2003 to correct PI field.) (Updated on 17-OCT-2003 to standardise
CC OS field)
XX
SQ Sequence 1631 BP, 554 A, 291 C, 339 G, 446 T, 0 U, 1 Other;

```

Alignment Scores:

```

Pred. No.: 1.5e-96 Length: 1631
Score: 1113.00 Matches: 223
Percent Similarity: 65.32% Conservative: 69
Best Local Similarity: 49.89% Mismatches: 121
Query Match: 46.57% Indels: 34

```


CC the incompatible interaction of plant gene expression relative to
 CC expression of the gene in an uninfected plant, in a mutant plant that
 CC does not express a gene associated with response to pathogenic infection,
 CC or in a corresponding incompatible or compatible interaction. (M1) is
 CC useful for conferring resistance to resistance or tolerance to a plant to
 CC bacterial, fungal or viral infection. The present sequence was used to
 CC illustrate the invention.

XX Sequence 1269 BP; 373 A; 285 C; 319 G; 292 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	Length:	Matches:	Mismatches:
Score: 5.01e-88	1269	193	59
Percent Similarity: 1022.50		Conservative: 59	109
Best Local Similarity: 69.23%		Mismatches: 109	
Query Match: 42.78%		Gaps: 3	
			2

US-10-691-374-2 (1-457) x ADA70539 (1-1269)

```

QY 75 ValIleasnValIleuSerPheGlyAlaLysGlyAspGlyThrTyraAspAnIleAla 94
Db 154 GTGTCGATGTGCGCAACATGCGCTTACGAGATGACACATGACACGAGGCA 213
QY 95 PheGluGlnAlaTPraAsnGlnAlaCysSerSerArgThrProValGlnPheValPro 114
Db 214 TTGCGAAAGGCAATGGGCTGCAAGCTTCCCTTCCTTCGCAACCTTCATTTCTCATCCCA 273
QY 115 LysAnIleuAsnTyLeuLeuLysGlnIleThrPheSerGlyProCysAsnSerSerIle 134
Db 274 AAGGCGAAGAGATACCTCACCACACATTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 333
QY 135 SerValIleuSerPheGlySerLeuGlnAlaSerSerIleSerAspTyLeuAspArg 154
Db 334 ACGTTCATGATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 393
QY 155 -----ArgLeuTPriIleAlaPheAspSerValGlnAsnLeuValAlaGlyGlyGly 172
Db 394 ACTATTAGGCACTGATATTATCTCAATGCTGATAGTGCGCTTATCTGCTGCTGCTGCTGCTG 453
QY 173 ThrIleasnGlyAsnGlyGlnValIleTPriProSerSerCysIleAsnIleuSerLeu 192
Db 454 ACTGTCATGATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 513
QY 193 ProCysArgAspAlaProThrAlaLeuThrPheThrAsnCysIleAsnIleuSerValAsn 212
Db 514 CCAATGCACTGATGCTCAACGCGCTTGAACATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 573
QY 213 AsnIleuSerIleAsnAlaGlnGlnIleIleIleIleIleIleIleIleIleIleIleIle 232
Db 574 AACTGAAAGCTACTAAACAGCCAAACAAATCCACATGCTGCTGCTGCTGCTGCTGCTGCTG 633
QY 233 ValAlaSerAsnIleuMetIleAsnAlaSerAlaLysSerProAsnThrAspGlyValHis 252
Db 634 AGGATCTTACGCTGCAATCACAGACACAGGACCTAGCCCAACACGCGCATTCAT 693
QY 253 ValSerAsnThrGlnTyLeuIleGlnIleSerAspThrIleIleGlyThrGlyAspAspCys 272
Db 694 ATCACAAAGAGTAAATATGATACAGTACAGGCTCATTAATCAAGACCGGAGATGACTGC 753
QY 273 IleSerIleValSerGlySerGlnAsnValGlnAlaThrAsnIleThrCysGlyProGly 292
Db 754 ATGTGATTTGAGATGAGTGAAGTAACTTACATGTCAGAAACATGCTGCTGCTGCTGCTGCTG 813
QY 293 HisGlyIleSerIleGlySerIleuGlySerGlyAsnSerGlyValAlaTyraValSerAsnVal 312
Db 814 CACGCGCATTCAGATCGGAGCTTACGCGATCAACAATCTGTAAGCTCATGCTCAACAAATGTC 873
QY 313 ThrValAsnGlnAlaLysIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 332
Db 874 ACCGTCGACACCGCTATATGCGCAACCAACGAGCTGCCATCAAGACATGCGAG 933
QY 333 GlyIleSerGlyGlnAlaSerAsnIleLysPheLeuAsnValGlnMetGlnAspValLys 352

```

```

Db 934 GCGGTTGGGGCTAGCGCAAGAAACATCGTGTCCAGAACATATATGAAATGTGG 993
QY 353 TyProIleIleIleIleAspIleAsnTyLeuAspArgValGluProCysIleGlnIlePhe 372
Db 994 AACCCATCATTTATTCACCAAGAACTAGTGTCTGCTACACCTGCAAGAAACAGACA 1053
QY 373 SerAlaValAlaGlnValLysAsnValAlaTyraGlnAsnIleLysGlyThrSerAlaThrTyrs 392
Db 1054 TCTGCGGTGCAGATGATGACATGTGCTTTCAGAAACATCAGGCGCAAGATGCTTCCAG 1113
QY 393 ValAlaIleLysPheAspCysSerThrAsnPheProCysGlyGlyIleIleMetGlnAsn 412
Db 1114 GAGGCTATCAACTGACCTGATGACAGAAATGACCTTGTCAAGAAATPACCTTGAAGGAC 1173
QY 413 IleAsnLeu---ValIleGlyLysSerGlyAspProSerGlyAlaThrCysLysAsnValHis 431
Db 1174 GTCAGCTCATCATGACAGGAGCTGTGATGATGTCAGAAAGCACTTGCAGAAACGCAAAA 1233
QY 432 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 433 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 434 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 435 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 436 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 437 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 438 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 439 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 440 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 441 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 442 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 443 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 444 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 445 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 446 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 447 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 448 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 449 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 450 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 451 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 452 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 453 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 454 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 455 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 456 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 457 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 458 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 459 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 460 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 461 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 462 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 463 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 464 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 465 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 466 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 467 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 468 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 469 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 470 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 471 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 472 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 473 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 474 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 475 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 476 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 477 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 478 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 479 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 480 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 481 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 482 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 483 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 484 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 485 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 486 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 487 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 488 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 489 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 490 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 491 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 492 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 493 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 494 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 495 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 496 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 497 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 498 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 499 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 500 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 501 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 502 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 503 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 504 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 505 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 506 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 507 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 508 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 509 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 510 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 511 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 512 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 513 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 514 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 515 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 516 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 517 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 518 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 519 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 520 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 521 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 522 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 523 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 524 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 525 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 526 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 527 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 528 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 529 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 530 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 531 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 532 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 533 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 534 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 535 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 536 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 537 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 538 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 539 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 540 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 541 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 542 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 543 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 544 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 545 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 546 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 547 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 548 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 549 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 550 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 551 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 552 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 553 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 554 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 555 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 556 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 557 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 558 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 559 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 560 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 561 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 562 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 563 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 564 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 565 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 566 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 567 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 568 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 569 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 570 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 571 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 572 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 573 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 574 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 575 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 576 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 577 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 578 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 579 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 580 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 581 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 582 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 583 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 584 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 585 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 586 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 587 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 588 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 589 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 590 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 591 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 592 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 593 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 594 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 595 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 596 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 597 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 598 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 599 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 600 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 601 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 602 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 603 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 604 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 605 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 606 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 607 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 608 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 609 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 610 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 611 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 612 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 613 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 614 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 615 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 616 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 617 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 618 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 619 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 620 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 621 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 622 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 623 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 624 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 625 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 626 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 627 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 628 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 629 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 630 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 631 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 632 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 633 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 634 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 635 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 636 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 637 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 638 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 639 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 640 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 641 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 642 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 643 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 644 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 645 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 646 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 647 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 648 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 649 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 650 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 651 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 652 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 653 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 654 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 655 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 656 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 657 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 658 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 659 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 660 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 661 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 662 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 663 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 664 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 665 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 666 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 667 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 668 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 669 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 670 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 671 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 672 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 673 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 674 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 675 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 676 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 677 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 678 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 679 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 680 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 681 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 682 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 683 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 684 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 685 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 686 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 687 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 688 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 689 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 690 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 691 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 692 PheAsnAsnAla 435
Db 1234 TCGAAGAAATCA 1245
QY 693 PheAsnAsnAla 435
Db 1234 TCGAAG
```

Alignment Scores:

Pred. No.:	5.75e-87	Length:	1182
Score:	1011.00	Matches:	187
Percent Similarity:	69.52%	Conservative:	57
Best Local Similarity:	53.28%	Mismatches:	105
Query Match:	42.30%	Indels:	2
DB:	8	Gaps:	1

US-10-691-374-2 (1-457) x ADA70540 (1-1182)

```

QY 72 GYILEYVAIIIEAENVALLIUSERPHEGIIYALYEGLYAEPQIYLYERTHYRARP 91
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 79 GGGAGTAATGTTTTCAGCATACAGAGCTACGGGGCTCAGAGAGGAGCATATGATGAC 138
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 92 AENIIIEAIPHEGLUGIINALATTPAENGIIUAICYSESERARGTHRPROVALGPNPE 111
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 139 ACCAAAGCATGGAGATACATGGGCTGCAGCTGCTCTCCAGCAAACTGCAGTTTGG 198
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 112 VALIAPROLYAENLYSANTYRLEUWLYEGIIIEIETHRPHSERGLYPROCYBARG 131
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 199 CTATATCCCAAGGAGCAAAATCCATGATCAAGACACAAACCTGCTGTCATGCAAA 258
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 132 SetSerIleSerValIyIlePheGlySerLeuGluAlaSerSerIyIleSerApyr 151
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 259 TCAGCATCTCATGTGATGAGTGAAGGTTGCTTGTCCAGAGAGGTCAAGCTGG 318
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 152 LyAaPaPyg-----ArgLeuTrpIIEAIPHEAPSERVALGIIAENILEUVALIGLY 169
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 319 ACDAAGAGAACCATAAAGCATGATTCGATCAGTGCATGCTGCTTACTGCTACT 378
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 170 GYIGLYGIIYTHRIIEANGLYANGIYGLINVALITPTTPROSERSERIYELIIEAN 189
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 379 GGTGGTGGAGCATATGATGAAATGCAAGATTGGTGCAAAATTCATGCAAAACCAAC 438
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 190 LydSerIleProCyAaPaPaIaProThraIaLeuThrPheThraPamCysIySaenLeu 209
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 439 TCAGAGCTTCATGCAAGAGCTCAAGGCACTGACATTCCTCCGAAAGATCTG 498
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 210 LyEValaEnaEnuLySerIySaenAlaGInGIIIEIIEIIEIIEIIEIIEIIEIIE 229
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 499 AAGGTGAGATATCTGAAGGTGGTGAACCCAGCAAAATCCAGATTTCAGTGGAGGTGTC 558
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 230 ThraEnValaIaIaSerEnuMetIIEANIIaSerIaIaIySerProAnThraP 249
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 559 ACCGATGATGATGCTGCTGCTGCTGATCAGCAACCAAGAAACACCCCAACACTGAT 618
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 250 GYIValaIEaIaSerEnThrGInTyRIIEGIIIEaSerApyrThRIIEGIIYTHRG 269
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 619 GGAATTCATATCACAGCAGCAGAGATGTTGAGGTACAGATCGCATGATCAAGACCGG 678
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 270 AAPAaPyCysIleSerIleValSerGlySerGlnAenValGlnAlaThraEnIIEThrCys 289
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 679 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 738
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 290 GYIYRPGIYHIIIEGLYIIESEIIEGLYSEIIEGLYSEIIEGLYSEIIEGLYSEIIE 309
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 739 GGAACGGGACAGCGGATCAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 798
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 310 SetAaEnValThraIaEnGluAlaIyIIEGLYIIEGLYIIEGLYIIEGLYIIEGLY 329
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 799 AACAAATGTCACCGTGCACAAATGCTGATGAGCAACGCAAGGAGCTGCAATCAAG 858
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 330 THRTGInGIIYGLYSEIIEGLYIIESEIIEGLYIIESEIIEGLYIIESEIIEGLY 349
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 859 ACATGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 918
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 350 AAPValIyEtyrProIIEIIEIIEIIEIIEIIEIIEIIEIIEIIEIIEIIEIIE 369
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 919 AATGTCGAAACCCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 978
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 370 GInGInPheSerIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 389
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 979 CAACGAAATCTGCAATGAGGTGAGCAATGCTGCTTCAAGAACATCAAGGGGCAAACT 1038

```

QY 390 AlathryIyValAlaIIEIIEIIEIIEIIEIIEIIEIIEIIEIIEIIEIIEIIE 409

DB 1039 GCATAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1098

QY 410 MetGluAaEnIIEaEnLeuValGIIYGLYSEIIEGLY 420

DB 1099 TTGGAGAAATGTCATATCTCATCTCTCAAGGAGAGA 1131

RESULT 15

ADA70633

ID ADA70633 standard; DNA; 1344 BP.

XX

AC ADA70633;

XX

DT 20-NOV-2003 (first entry)

XX

DE Rice gene, SEQ ID 3956.

XX

KM Plant; bacterial infection; fungal infection; viral infection; rice;

XX

KM gene; ds.

XX

OS Oryza sativa.

XX

PN MO2003000898-A1.

XX

PD 03-JAN-2003.

XX

PF 22-JUN-2001; 2001WO-1B001105.

XX

PR 22-JUN-2001; 2001WO-1B001105.

XX

PA (SYN) SYNGENTA PARTICIPATIONS AG.

XX

PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;

XX

PI Katagiri F, Quan S, Tao Y, Whitcham S, Xie Z, Zhu T, Zou G;

XX

DR WPI; 2003-175290/17.

XX

PT Identifying at least one gene involved in plant resistance or response to

XX

PT pathogenic infection for conferring resistance or tolerance to a plant to

XX

PT bacterial, fungal or viral infection by determining or detecting plant

XX

PT gene expression.

XX

PS Claim 6; SEQ ID NO 3956; 899bp; English.

XX

CC The present invention relates to a method (M1) for identifying genes

XX

CC involved in plant resistance or response to pathogenic infection. M1

XX

CC comprises identifying a gene whose expression is significantly altered in

XX

CC the incompatible interaction of plant gene expression relative to

XX

CC expression of the gene in an uninfected plant, in a mutant plant that

XX

CC does not express a gene associated with response to pathogenic infection,

XX

CC or in a corresponding incompatible or compatible interaction. (M1) is

XX

CC useful for conferring resistance to resistance or tolerance to a plant to

XX

CC bacterial, fungal or viral infection. The present sequence was used to

XX

CC illustrate the invention.

XX

SQ Sequence 1344 BP; 384 A; 298 C; 331 G; 331 T; 0 U; 0 Other;

XX

Alignment Scores:

Pred. No.:	9.7e-86	Length:	1344
Score:	999.00 <td>Matches:</td> <td>193</td>	Matches:	193
Percent Similarity:	64.23% <td>Conservative:</td> <td>59</td>	Conservative:	59
Best Local Similarity:	49.23% <td>Mismatches:</td> <td>118</td>	Mismatches:	118
Query Match:	41.80% <td>Indels:</td> <td>22</td>	Indels:	22
DB:	8	Gaps:	4

US-10-691-374-2 (1-457) x ADA70633 (1-1344)

```

QY 75 ValIIEaEnValIIEuSerPheGIIYALYEGLYAEPQIYLYERTHYRARPENIIIEA 94
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 154 GTGTGATGATGTCGGAACATAGCTCTTACGAGATGACAAACATGATGATCAAAAGCA 213

```

```
OY 95 PheGluGlnAlaTrpAsnGlnAlaCysSerSerArgThrProValGlnPheValValPro 114
Db 214 TTGTCAAAAGGCATGGGCGCGCTGTGCTCTCTTTGCAACCTTCATTTGTTCTGTTCCA 273
OY 115 LysAsnLysAsnTrpLeuLeuLysGlnIleThrPheSerGlyProCysArgSerSerIle 134
Db 274 AAGGCGCAAGATACCTCACCAAGACACATTTACATTTGGCCCAATGCAATCCAGCATC 333
OY 135 SerValLysIlePheGlySerLeuGlnAlaSerSerLysIleSerAspTrpLysAspArg 154
Db 334 ACGTTCATGATAGAGGGTACTTTGGTGGCTCTCCCAAGAGGTGAGATTGGACAGAGAA 393
OY 155 -----ArgLeuTrpIleAlaPheAspSerValGlnAsnLeuValValGlyGlyGly 172
Db 394 ACTATTAGCACTGGATTATGTTCAATGATGTAATGGTCTTACTGTGCGGTGG 453
OY 173 ThrIleAsnGlyAsnGlyGlnValTrpTrpProSerSerCysLysIleAsnLysSerLeu 192
Db 454 ACGGTCGATGGAATGGCAAGATTGGTGCAAAATCTTGCAAAACGAATGCGAAACTT 513
OY 193 ProCysArgAspAlaProThrAlaLeuThrPheTrpAsnCysLysAsnLeuLysValAsn 212
Db 514 GCATGCACTGATATCTCCACGGCCTTGACATTTCACTCCCTTCAAAATCGAAAGTGGAG 573
OY 213 AsnLeuLysSerLysAsnAlaGlnGlnIleHisIleLysPheGlySerCysThrAsnVal 232
Db 574 AACTTGAAGCTACTAAACAGCCACAAATCCACATGTCAGTGAGATTCACAAATGTT 633
OY 233 ValIleSerAsnLeuMetIleAsnAlaSerAlaLysSerProAsnThrAspGlyValHis 252
Db 634 AGCATCTCTGGCCGACATCACAGCACAGGCACTAGCCCAACACCGATGGCATCCAT 693
OY 253 ValSerAsnThrGlnTrpIleGlnIleSerAspThrIleIleGlyThrGlyAspAspCys 272
Db 694 ATCCACAGAGTAAATAATGTACAAAGTACAGGCTGCACAACTCAAGACCGGGACGATGC 753
OY 273 IleSerIleValSerGlySerGlnAsnValGlnAlaThrAsnIleThrCysGlyProGly 292
Db 754 ATGTGATCGAGATGAGATGGAACCTTCATGTCAGAAACATGGTGGCGGACCGAGA 813
OY 293 HisGlyIleSerIleGlySerLeuGlySerGlyAsnSerGlnAlaTrpValSerAsnVal 312
Db 814 CACGGCATCGCATTCGGGACCTTAGCGATCACACTCTGAGGCTCATGTCAACAAATGTC 873
OY 313 ThrValAsnGlnAlaLysIleIleGlyAlaGluAsnGlyValArgIleLysThrTrpGln 332
Db 874 ACCATCGGCACTGTCAAGCTATATGGCACACCAACGAGCTCGCATCAAGACATGGCAG 933
OY 333 GlyGlySerGlyGlnAlaSerAsnIleLysPheLeuAsnValGluMetGlnAspValLys 352
Db 934 GGTGGTCGGGGGTACCGGAAGTACATGTTTCCAGAACATGATCATGGAAATGTTGG 993
OY 353 TyrProIleIleIleAspGlnAsnTrpCysAspArgValGluProCys----- 368
Db 994 AACCCGTCATTATGATCAAAACTACTGTGACTGTGCAACACCATGCAAGAAACAGTTG 1053
OY 369 -----IleGlnGlnPheSerAla 374
Db 1054 TCCATTGTGTGAATAATATGGAAGCTTAATTAATTAGTGTGTTGTTCTCAAGCATCTGCA 1113
OY 375 ValGlnValLysAsnValValTrpGluAsnIleLysGlyThrSerAlaThrLysValAla 394
Db 1114 GTTTCAGATTAGCAATGTGTCTTCAAGAACATCAGGGGCAACAGTCTTCCAAAGGAGGCT 1173
OY 395 IleLysPheAspCysSerThrAsnPheProCysGluGlyIleIleMetGluAsnIleAsn 414
Db 1174 ATCAAGCTGAGCTGACAGAAATGATCTTGCCAAAGCATTAACCTTAATGATGTCAAG 1233
OY 415 Leu---ValGlyGluSerGlyLysProSerGlnAlaThrCysLysAsnValHisPheAsn 433
Db 1234 CTCACCTGCACAGGAGGTGGTGTGATGCGAAGACACTTCCGAAACGCAAAATGGAAG 1293
OY 434 AsnAlaGluHisValThrProHis---CysThrSer 444
```

Db 1294 AATCAAGGACAGTTGTTCACAGCCTTGTGCTTC 1329

Search completed: March 4, 2005, 22:54:05
Job time : 673 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 4, 2005, 19:53:34 ; Search time 3526 Seconds
(without alignments)
4933.457 Million cell updates/sec

Title: US-10-691-374-2
Perfect score: 2390
Sequence: 1 NTGKNSILLIIFASSIS.....VTPHCTSLSEIDELLYNY 457

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues
Total number of hits satisfying chosen parameters: 6847908

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+.p2n.model -DEV=xlp
-O=/cgn2.1/USPTO.spool.p/US10691374/runat.28022005.120706.20996/app_query.fasta_1.647
-DB=EST -QFMT=fastlap -SUFFIX=rst -MIMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blonum62 -TRANS=human40.cdi -LIST=45
-NOALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pro -NOM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10691374 -@CGN 1 1 5180 -runat.28022005.120706.20996 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WMRN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
EST.*
1: gb_esc1:*
2: gb_esc2:*
3: gb_hic:*
4: gb_esc3:*
5: gb_esc4:*
6: gb_esc5:*
7: gb_esc6:*
8: gb_gsa1:*
9: gb_gsa2:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1347	56.4	797	4	BM410796 EST585123
2	1343	56.2	786	4	BM412817 EST587144
3	1331	55.3	774	4	BM413004 EST587331
4	1292	54.1	767	4	BM412528 EST586855
5	1282	53.6	742	4	BM413282 EST587609
6	1278	53.5	749	4	BM413343 EST587670
7	1250	51.0	709	2	BM413343 EST311335
8	1212	50.7	696	2	AM442253 EST311649
9	1192	49.9	687	2	AM442335 EST311731

10	1168	48.9	677	2	AM442241	AM442241 EST311637
11	1160	48.5	675	4	BM408546	BM408546 EST582873
12	1157.5	48.4	672	2	AM221785	AM221785 EST298596
13	1128	47.2	643	2	AM223400	AM223400 EST300211
14	1118	46.8	672	2	AM222222	AM222222 EST299033
15	1110	46.4	703	4	BM536089	BM536089 EST589111
16	1109	46.4	728	4	BM413075	BM413075 EST587402
17	1107	46.3	632	2	BE434493	BE434493 EST405571
18	1107	46.3	632	2	AM442280	AM442280 EST311676
19	1106	46.3	632	2	BE432845	BE432845 EST399470
20	1102	46.1	632	4	BM409252	BM409252 EST583579
21	1099	46.0	630	2	BE431613	BE431613 EST336428
22	1097	45.9	632	4	BM410538	BM410538 EST584865
23	1094	45.8	1498	3	CNSDA7SM	EX822965 Arabidops
24	1077	45.1	617	2	BE460938	BE460938 EST412357
25	1072	44.9	619	4	BM413202	BM413202 EST87529
26	1058	44.3	614	2	BE433180	BE433180 EST39709
27	1055	44.1	670	4	BM411547	BM411547 EST585874
28	1050	43.9	604	2	BE461204	BE461204 EST412623
29	1041	43.6	601	2	AM222835	AM222835 EST299646
30	1032	43.2	598	2	AM223911	AM223911 EST300722
31	1016	42.5	585	2	BE436917	BE436917 EST408035
32	1014	42.4	593	2	AM222726	AM222726 EST299537
33	1013.5	42.4	588	2	BE432533	BE432533 EST399062
34	1010	42.3	604	2	BE433321	BE433321 EST399850
35	1003	42.0	1182	9	CL972417	CL972417 OeIFCC022
36	1001	41.9	576	2	AM222864	AM222864 EST299675
37	997	41.7	1248	9	CL972416	CL972416 OeIFCC022
38	994.5	41.6	1344	9	CL972420	CL972420 OeIFCC022
39	994	41.6	572	2	AM222731	AM222731 EST299542
40	990	41.4	567	2	AM441297	AM441297 EST311593
41	989	41.4	563	2	AM441402	AM441402 EST310798
42	988	41.3	565	2	BE460910	BE460910 EST412329
43	987	41.3	567	2	AM223661	AM223661 EST300472
44	984	41.2	561	2	BE437064	BE437064 EST408182
45	981	41.0	560	2	BE432446	BE432446 EST398975

ALIGNMENTS

RESULT 1
BM410796
LOCUS
DEFINITION
BMA10796
797 bp mRNA linear EST 22-JUN-2002
CLES54L15 5' end, mRNA sequence.

ACCESSION
BM410796
VERSION
BM410796.1 GI:18262426
KEYWORDS
SOURCE
ORGANISM
Lycopersicon esculentum (tomato)

REFERENCE
Alcala,J., Vrebalov,J., White,R., Viston,T., Karamycheva,S.A.,
Tsai,J., Bougri,O., Kirnase,E., Utechback,T., Van Aken,S.,
Roming,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and
Giovannoni,J.
Generation of ESTs from tomato fruit tissue, breaker stage (2002)
Unpublished (2002)
Contact: CUGI

TITLE
JOURNAL
COMMENT
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>
This clone is available through the Clemson University Genomics
Institute
Contact: CUGI

FEATURES
source
Seq primer: T3
Location/Qualifiers
1..797
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"

```

/db_xref="taxon:4081"
/clone="cLBG54L15"
/issue_type="Pericarp"
/dev_stage="breaker"
/lab_host="SOLR"
/clone_1lb="tomato breaker fruit"
/note="Vector: pBluescriptSkmCudapt; Site_1: EcoRI;
Site_2: XhoI; supplier: Boyce Thompson Institute;
sequencing: The Institute for Genomic Research. Fruit
were harvested at the breaker stage (first sign of
lycopene accumulation on the blossom end of fruit). Fruit
were cut in half and the seeds and locules were discarded
prior to freezing the pericarp."

```

ORIGIN

Alignment Scores:

```

Pred. No.: 1.57e-144 Length: 797
Score: 1347.00 Matches: 261
Percent Similarity: 98.50% Conservative: 1
Best Local Similarity: 98.12% Mismatches: 3
Query Match: 56.36% Indels: 1
DB: 4 Gaps: 0

```

US-10-691-374-2 (1-457) x BM410796 (1-797)

```

QY 104 SerSerArgThrProValGlnPheValValProLysAsnLysAsnTyrlLeuLeuLysGln 123
DB 2 TCATCTAGAACACCGTTCAATTGTGGTCTTAAACAGATTATCTTCAGACCA 61
QY 124 lLeThrPheSerGlyProCysArgSerSerLysValLysIlePheGlySerLeuGlu 143
DB 62 ATCACTTTCAAGTCCAGTCAAGATCTTAAATTCAGTAAAGATTTTGGATCTTAGAA 121
QY 144 AlAsSerLysIleSerAspTyrlLysAspArgLeuTrpIleAlaPheAspSerVal 163
DB 122 GCATCTAGTAAATTTGACACTACAAAGATGAAAGCTTGCTTTGATAGTGT 181
QY 164 GlnAsnLeuValValGlyGlyGlyGlyThrlIleAsnGlyAsnGlyGlnValTrpTrpPro 183
DB 182 CAAATTTAGTTGTTGAGAGAGAGAACTATCAATGCGCAATGCAATGCTGCGCA 241
QY 184 SerSerCysAlaLysIleAsnLysSerLeuProCysArgAspAlaTrpThrAlaLeuThrPhe 203
DB 242 AGTCTTGGCAAAATTAATTAATCACTGCCATGACGAGATGCAAGGCTTTAACTTTC 301
QY 204 TrpAsnCysLysAsnLeuLysValAsnAsnLeuLysSerLysAsnAlaGlnGlnIleHis 223
DB 302 TGGAAATTCGCAAAATTTGAAAGGATTAATCTAAAGATGAAATGCAACAAATTCAT 361
QY 224 lLeLysPheGlnSerCysThrAsnValValAlaSerAsnLeuMetIleAsnAlaSerAla 243
DB 362 ATCAATTTGAGTCAATGCACTAATGTTGAGCTTCAAATTTGATATCAATGCTTCAGCA 421
QY 244 LysSerProAsnThrAspGlyValHisValSerAsnTrpGlnTyrlIleGlnIleSerAsp 263
DB 422 AAGAGCGCAAAATCAATGATGAGTCAATGATCAATATCAATATATTCATATTCAT 481
QY 264 ThrIleIleGlyThrlLysAspAspCysIleSerIleValLysSerGlySerGlnAsnValGln 283
DB 482 ACTATTTAGGAACGGTATGATGTTTTCATTAATTTGTTTGGATTCGCAAAATGTCAG 541
QY 284 AlAtThrAsnIleThrCysGlyProGlyHisGlyIleSerIleGlySerLeuGlySerGly 303
DB 542 GCGCAAAATTAATTAATCTTGGTCCAGGTGATGATTAATGATGAACTTGAGATCTGGA 601
QY 304 AsnSerGlnAlaTyrlValSerAsnValThrValAsnGlnAlaLysIleIleGlyAlaGlu 323
DB 602 AATTCAGAGCTTATGTTCTTAATGTTACTTAATATGAAGCCAAATTAATCGGTGCGAA 661
QY 324 AsnGlyValAlaGlyIleLysThrTrpGlnGlyGlySerGlyGlnAlaSerAsnIleLysPhe 343
DB 662 AATGAGTTAGATCAAGACTTGCGAGGAGATCTGAGCAAGTACCACTCACTCAATTAAT 721

```

```

QY 344 LeuAsnValAlaLysGlnAspValLysTyrlProIleIleAlaAspGlnAsnTyrlCys-As 363
DB 722 CTGAATGTGAAATGCAAGCGTTTGTATCCATATTAATGACCAAACTATTTGGTCA 781
QY 363 PARValGluProCys 368
DB 782 TCGAGTTTGACCATGT 797

```

RESULT 2

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

```

TITLE
JOURNAL
COMMENT
Clemson University Genomics Institute
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
This clone is available through the Clemson University Genomics
Institute
Seq primer: T3.

```

FEATURES

source

```

1..786
location/Qualifiers

```

```

/molecule="mRNA"
/organism="Lycopersicon esculentum"
/cultivar="7496"
/db_xref="taxon:4081"
/clone="cLBG61A20"
/tissue_type="Pericarp"
/dev_stage="breaker"
/lab_host="SOLR"
/clone_1lb="tomato breaker fruit"
/note="Vector: pBluescriptSkmCudapt; Site_1: EcoRI;
Site_2: XhoI; supplier: Boyce Thompson Institute;
sequencing: The Institute for Genomic Research. Fruit
were harvested at the breaker stage (first sign of
lycopene accumulation on the blossom end of fruit). Fruit
were cut in half and the seeds and locules were discarded
prior to freezing the pericarp."

```

ORIGIN

Alignment Scores:

```

Pred. No.: 4.46e-144 Length: 786
Score: 1343.00 Matches: 257
Percent Similarity: 98.09% Conservative: 0
Best Local Similarity: 98.09% Mismatches: 5
Query Match: 56.19% Indels: 0
DB: 4 Gaps: 0

```

US-10-691-374-2 (1-457) x BM412817 (1-786)

```

QY 95 PheGluGlnAlaTrpPheGlnAlaCysSerSerArgThrProValGlnPheValPro 114
DB 1 TTTCGCAAGGATGAGATGAAAGCATGTTCACTAGAACACCTGTTCAATTTGTGTTCT 60
QY 115 LysAsnLysAsnTyrlLeuLeuLysGlnIleThrPheSerGlyProCysArgSerSerIle 134
DB 61 AAAAAAGATTTATCTTCTCAAGCAATCACTTTTCAGTCCATGACATCTTCTATT 120

```

QY 135 SerVallysllephleglySerleuglnalaserSerlylleSeraspTyllyasparg 154
DB 121 TCAGTAAAGATTTTGGATCCCTTAGACACTTAAATTTTCAGACTTCACAAATAGA 180
QY 155 ArgLeuTPPllleAlaPheaspSerValGlnleuValValGlyGlyGlyThrIle 174
DB 181 AGGCTTGATGCTTCTTGATGATGCTTCAAAATTTAGTTGTTGGAGAGAGAACTATC 240
QY 175 AenglyAenGlyGlnValTPPTrpProSerSerCyslylleAenlyleSerleuProCys 194
DB 241 AATGCATGAGCAAGATATGATGGCCAAAGTTCTTCAAAATTAATCACTGCATGCG 300
QY 195 ArgAenAlaProThrAlaLeuThrPheThrPheCyslylleAenleuValAenleu 214
DB 301 AGGGATGACCAACCGCTTAACCTTCTGGAATTCGAAAAATTTGAAAGTGAATATCTA 360
QY 215 LysSerlyAenAlaGlnGlnIleHleIlelylePheGlySerCysThrAenValAla 234
DB 361 AAGGTAAAGATGACCAACCAATTCATCAATTTAGTCACTGATGATGTTAGCT 420
QY 235 SerAenleuMelileAenAlaserAlaLysSerProAenThrAspGlyValHleValSer 254
DB 421 TCAATTTGATGATCAATGCTTCACGCAAGAGCCAAATGATGATGATGATGATCA 480
QY 255 AenThrGlnTylleGlnIleSerAenThrIleIleGlyThrGlyAspAspCysIleSer 274
DB 481 AATCTCATATATATCAATATCTGATCTATTTAGGAACGATGATGATGATTTCA 540
QY 275 IleValSerGlySerGlnAenValGlnAlaThrAenIleThrCysGlyProGlyHleGly 294
DB 541 ATTGTTCTGATTCGAAATGAGGAGGACCAAAATTTACTTGTTGTTGTTGTTGTTGTT 600
QY 295 IleSerIleGlySerleuglySerGlyAenSerGlnAlaTyrValSerAenValThrVal 314
DB 601 ATAAGTATTTGGAGCTTAGGATCTGAAATTCAGAAAGCTATGCTTAATGTTCTGTA 660
QY 315 AenglnAlaLysIleIleGlyAlaGlnAenglyValAenglyIleYthrTrpGlnGly 334
DB 661 AATGAAGCCAAATATTCGTCGCCAAATAGAGTTAGATCAACACTTGCGAGGGAGA 720
QY 335 SerGlyGlnAlaserAenIlelylePheAenValGlnMetGlnAspValIleYthrPro 354
DB 721 TCTGGAACAGCTGACACATCAATTTCTGAATGTGAAATGCAAGACCTTAATATCCC 780
QY 355 IleIle 356
DB 781 ATTAAT 786
RESULT 3
LOCUS BM413004 774 bp mRNA linear EST 22-JAN-2002
DEFINITION EST587331 tomato breaker fruit Lycopersicon esculentum cDNA clone
c18661n6.5' end, mRNA sequence.
ACCESSION BM413004
VERSION BM413004.1 GI:18264634
KEYWORDS EST.
SOURCE Lycopersicon esculentum (tomato)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; Lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 774)
REFERENCE
AUTHORS Alcala,J., Vrebalov,J., White,R., Vision,T., Karamycheva,S.A.,
Teal,J., Bougri,O., Kirness,B., Utecherback,T., Van Aken,S.,
Roming,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and
Giovannoni,J.
TITLE Generation of ESTs from tomato fruit tissue, breaker stage (2002)
JOURNAL Unpublished (2002)
CONTACT: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA

Email: <http://www.genome.clemson.edu/orders/index.html>
This clone is available through the Clemson University Genomics
Institute
Seq primer: T3.
FEATURES
source
Location/Qualifiers
1..774
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="c18661n6"
/tissue_type="pericarp"
/dev_stage="breaker"
/lab_host="SOUR"
/clone_id="tomato breaker fruit"
/note="Vector: pBluescriptSMC2adp; Site 1: EcoRI;
Site 2: XhoI; supplier: Boyce Thompson Institute;
sequencing: The Institute for Genomic Research. Fruit
were harvested at the breaker stage (first sign of
lycopene accumulation on the blossom end of fruit). Fruit
were cut in half and the seeds and locules were discarded
prior to freezing the pericarp."
ALIGNMENT SCORES:
Pred. No.: 1,52e-141 Length: 774
Score: 1321.00 Matches: 254
Percent Similarity: 98.84% Conservative: 1
Best Local Similarity: 98.45% Mismatches: 2
Query Match: 55.27% Indels: 1
DB: 4 Gaps: 0
US-10-691-374-2 (1-457) x BM413004 (1-774)
QY 88 LysThrTyrAenAlaPheGlnGlnAlaTPAenGlnAlaCysSerSerAenThr 107
DB 3 AAAACATATGATTAATTTGATTTGACGACATGAAATGAACATGTTCACTAGAAC 62
QY 108 ProValGlnPheValAlaProlyleAenlyleAenlyleleuLeuGlnIleThrPheSer 127
DB 63 CCGTTCATATTTGCTGCTTCAAAAACAAATTTCTTCTTCACCAATATCACCTTTTCA 122
QY 128 GlyProCysArgSerSerIleSerValIlePheGlySerleuglnAlaserSerlyls 147
DB 123 GGTCCATCAGATCTTCTTAATTTCACTAAAGATTTTGGATCTTCAAGCATCTTGTAA 182
QY 148 IleSerAenPyllyleAenPargArgLeuTPPllleAlaPheaspSerValGlnAenleuVal 167
DB 183 ATTTCAGACTACAAAGATAGAAAGCTTTGATGCTTTTGAATGATGATGATGATGAT 242
QY 168 ValGlyGlyGlyIleThrIleAenglyAenglyGlnValTPPTrpProSerSerCyslyls 187
DB 243 GTTGAGAGAGAGGAATCATCATGATGACAAATGATGCTGCGCAAGTTCTTGCAAA 302
QY 188 IleAenlyleSerleuProCysArgAspAlaProThrAlaLeuThrPheThrAenCyslyls 207
DB 303 ATTAATTAATATCTGCGCATGACGAGATCACCAACGCTTAACCTTCTGGAATGCAAA 362
QY 208 AenleuValAenAenleuLySerlyleAenAlaGlnGlnIleHleIlelylePheGln 227
DB 363 AATTTGAAAGTCAATATCTTAAGATGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 422
QY 228 SerCysThrAenValAlaLaserAenleuValProThrAlaLeuThrPheThrAenCyslyls 247
DB 423 TCATGACATTAATGTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 482
QY 248 ThrAenGlyValIleValSerAenThrGlnTyrIleGlnIleSerAenThrIleIleGly 267
DB 483 ACTGATGAGTCCATGATCAATATCTCAATATATTTCAATATTTCAATATTTCAATATTT 542
QY 268 ThrGlyAenAspCysIleSerIleValSerGlnAenValGlnAlaThrAenIle 287
DB 543 ACAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 602

Oy	288	ThrcysgslpProGlyVhIsgIlylEsErtIlgIySerIenuIySergIyansergIula	307		
Db	603	ACTGTGATCCAGGATCATGTGATATAGTATGGAAGCTTAGATCTGGAAATTCAGAGCT	662		
Oy	308	TyValSerAenValTrValAenGuaIalysErtIlgIyAlaGuaenGlyValArg	327		
Db	663	TATGTGCTATGTTACTGTAAATGAAAGCCAAATATTCGGTGGCGAAATGAGTTAGG	722		
Oy	328	IlElyThr-TTrpGlnGlyGlySergIyGlnalaseAenIlElySPhenLeu	344		
Db	723	ATCAAGACTTGGCAGGAGGATCTGCAACAGTACCAATCAATTCTTG	774		
RESULT 4	BM412528	767 bp	mRNA	linear	EST 22-JAN-2002
LOCUS	BM412528				
DEFINITION	EST56885 tomato breaker fruit lycopersicon esculentum cDNA clone				
ACCESSION	C1EG6014 5' end, mRNA sequence.				
VERSION	BM412528				
KEYWORDS	BM412528.1 GI:18264158				
SOURCE	EST.				
ORGANISM	Lycopersicon esculentum (tomato)				
REFERENCE	Lycopersicon esculentum				
AUTHORS	Bukaryoka, VitiDiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon. 1 (bases 1 to 767)				
	Alcala,J., Vrebalov,J., White,R., Vision,T., Karamycheva,S.A., Tsal,J., Bougri,O., Kirpness,E., Utecherback,T., Van Aken,S., Roming,C.M., Frazer,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.				
	Generation of ESTs from tomato fruit tissue, breaker stage (2002)				
TITLE	Unpublished (2002)				
JOURNAL	Contact: CUGI				
COMMENT	Clemson University Genomics Institute				
	100 Jordan Hall, Clemson, SC 29634, USA				
	Email: http://www.genome.clemson.edu/orders/index.html				
	This clone is available through the Clemson University Genomics Institute				
FEATURES	Seq primer: T3.				
source	Location/Qualifiers				
	1..767				
	/organism="Lycopersicon esculentum"				
	/mol_type="mRNA"				
	/cultivar="TA496"				
	/db_xref="taxon:4081"				
	/clone="C1EG6014"				
	/tissue_type="Pericarp"				
	/dev_stage="Breaker"				
	/lab_host="SOLR"				
	/clone_lib="tomato breaker fruit"				
	/note="Vector: pBluescriptSkmDap; Site:1: EcoRI, Site:2: XhoI; supplier: Boyce Thompson Institute; sequenced: The Institute for Genomic Research. Fruit were harvested at the breaker stage (first sign of lycopene accumulation on the blossom end of fruit). Fruit were cut in half and the seeds and locules were discarded prior to freezing the pericarp."				
ORIGIN					
Alignment Scores:					
Pred. No.:	3,39e-138	Length:	767		
Score:	1292.00	Matches:	248		
Percent Similarity:	100.00%	Conservative:	0		
Best Local Similarity:	100.00%	Mismatches:	0		
Query Match:	54.06%	Indels:	0		
DB:	4	Gaps:	0		
US-10-631-374-2 (1-457) x BM412528 (1-767)					
Oy	1 MetValIleGlnArgAnSerIleLeuLeuIleIleIlePheAlaSerSerIleSer				20

Db	Accession	Gene Name	Species	Length (bp)	Position (nt)	Sequence	GC Content (%)	GC Skew	GC Bias	GC Bias2	GC Bias3	GC Bias4	GC Bias5	GC Bias6	GC Bias7	GC Bias8	GC Bias9	GC Bias10	GC Bias11	GC Bias12	GC Bias13	GC Bias14	GC Bias15	GC Bias16	GC Bias17	GC Bias18	GC Bias19	GC Bias20	GC Bias21	GC Bias22	GC Bias23	GC Bias24	GC Bias25	GC Bias26	GC Bias27	GC Bias28	GC Bias29	GC Bias30	GC Bias31	GC Bias32	GC Bias33	GC Bias34	GC Bias35	GC Bias36	GC Bias37	GC Bias38	GC Bias39	GC Bias40	GC Bias41	GC Bias42	GC Bias43	GC Bias44	GC Bias45	GC Bias46	GC Bias47	GC Bias48	GC Bias49	GC Bias50	GC Bias51	GC Bias52	GC Bias53	GC Bias54	GC Bias55	GC Bias56	GC Bias57	GC Bias58	GC Bias59	GC Bias60	GC Bias61	GC Bias62	GC Bias63	GC Bias64	GC Bias65	GC Bias66	GC Bias67	GC Bias68	GC Bias69	GC Bias70	GC Bias71	GC Bias72	GC Bias73	GC Bias74	GC Bias75	GC Bias76	GC Bias77	GC Bias78	GC Bias79	GC Bias80	GC Bias81	GC Bias82	GC Bias83	GC Bias84	GC Bias85	GC Bias86	GC Bias87	GC Bias88	GC Bias89	GC Bias90	GC Bias91	GC Bias92	GC Bias93	GC Bias94	GC Bias95	GC Bias96	GC Bias97	GC Bias98	GC Bias99	GC Bias100	GC Bias101	GC Bias102	GC Bias103	GC Bias104	GC Bias105	GC Bias106	GC Bias107	GC Bias108	GC Bias109	GC Bias110	GC Bias111	GC Bias112	GC Bias113	GC Bias114	GC Bias115	GC Bias116	GC Bias117	GC Bias118	GC Bias119	GC Bias120	GC Bias121	GC Bias122	GC Bias123	GC Bias124	GC Bias125	GC Bias126	GC Bias127	GC Bias128	GC Bias129	GC Bias130	GC Bias131	GC Bias132	GC Bias133	GC Bias134	GC Bias135	GC Bias136	GC Bias137	GC Bias138	GC Bias139	GC Bias140	GC Bias141	GC Bias142	GC Bias143	GC Bias144	GC Bias145	GC Bias146	GC Bias147	GC Bias148	GC Bias149	GC Bias150	GC Bias151	GC Bias152	GC Bias153	GC Bias154	GC Bias155	GC Bias156	GC Bias157	GC Bias158	GC Bias159	GC Bias160	GC Bias161	GC Bias162	GC Bias163	GC Bias164	GC Bias165	GC Bias166	GC Bias167	GC Bias168	GC Bias169	GC Bias170	GC Bias171	GC Bias172	GC Bias173	GC Bias174	GC Bias175	GC Bias176	GC Bias177	GC Bias178	GC Bias179	GC Bias180	GC Bias181	GC Bias182	GC Bias183	GC Bias184	GC Bias185	GC Bias186	GC Bias187	GC Bias188	GC Bias189	GC Bias190	GC Bias191	GC Bias192	GC Bias193	GC Bias194	GC Bias195	GC Bias196	GC Bias197	GC Bias198	GC Bias199	GC Bias200	GC Bias201	GC Bias202	GC Bias203	GC Bias204	GC Bias205	GC Bias206	GC Bias207	GC Bias208	GC Bias209	GC Bias210	GC Bias211	GC Bias212	GC Bias213	GC Bias214	GC Bias215	GC Bias216	GC Bias217	GC Bias218	GC Bias219	GC Bias220	GC Bias221	GC Bias222	GC Bias223	GC Bias224	GC Bias225	GC Bias226	GC Bias227	GC Bias228	GC Bias229	GC Bias230	GC Bias231	GC Bias232	GC Bias233	GC Bias234	GC Bias235	GC Bias236	GC Bias237	GC Bias238	GC Bias239	GC Bias240	GC Bias241	GC Bias242	GC Bias243	GC Bias244	GC Bias245	GC Bias246	GC Bias247	GC Bias248	GC Bias249	GC Bias250	GC Bias251	GC Bias252	GC Bias253	GC Bias254	GC Bias255	GC Bias256	GC Bias257	GC Bias258	GC Bias259	GC Bias260	GC Bias261	GC Bias262	GC Bias263	GC Bias264	GC Bias265	GC Bias266	GC Bias267	GC Bias268	GC Bias269	GC Bias270	GC Bias271	GC Bias272	GC Bias273	GC Bias274	GC Bias275	GC Bias276	GC Bias277	GC Bias278	GC Bias279	GC Bias280	GC Bias281	GC Bias282	GC Bias283	GC Bias284	GC Bias285	GC Bias286	GC Bias287	GC Bias288	GC Bias289	GC Bias290	GC Bias291	GC Bias292	GC Bias293	GC Bias294	GC Bias295	GC Bias296	GC Bias297	GC Bias298	GC Bias299	GC Bias300	GC Bias301	GC Bias302	GC Bias303	GC Bias304	GC Bias305	GC Bias306	GC Bias307	GC Bias308	GC Bias309	GC Bias310	GC Bias311	GC Bias312	GC Bias313	GC Bias314	GC Bias315	GC Bias316	GC Bias317	GC Bias318	GC Bias319	GC Bias320	GC Bias321	GC Bias322	GC Bias323	GC Bias324	GC Bias325	GC Bias326	GC Bias327	GC Bias328	GC Bias329	GC Bias330	GC Bias331	GC Bias332	GC Bias333	GC Bias334	GC Bias335	GC Bias336	GC Bias337	GC Bias338	GC Bias339	GC Bias340	GC Bias341	GC Bias342	GC Bias343	GC Bias344	GC Bias345	GC Bias346	GC Bias347	GC Bias348	GC Bias349	GC Bias350	GC Bias351	GC Bias352	GC Bias353	GC Bias354	GC Bias355	GC Bias356	GC Bias357	GC Bias358	GC Bias359	GC Bias360	GC Bias361	GC Bias362	GC Bias363	GC Bias364	GC Bias365	GC Bias366	GC Bias367	GC Bias368	GC Bias369	GC Bias370	GC Bias371	GC Bias372	GC Bias373	GC Bias374	GC Bias375	GC Bias376	GC Bias377	GC Bias378	GC Bias379	GC Bias380	GC Bias381	GC Bias382	GC Bias383	GC Bias384	GC Bias385	GC Bias386	GC Bias387	GC Bias388	GC Bias389	GC Bias390	GC Bias391	GC Bias392	GC Bias393	GC Bias394	GC Bias395	GC Bias396	GC Bias397	GC Bias398	GC Bias399	GC Bias400	GC Bias401	GC Bias402	GC Bias403	GC Bias404	GC Bias405	GC Bias406	GC Bias407	GC Bias408	GC Bias409	GC Bias410	GC Bias411	GC Bias412	GC Bias413
----	-----------	-----------	---------	-------------	---------------	----------	----------------	---------	---------	----------	----------	----------	----------	----------	----------	----------	----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------	------------

100 Jordan Hall, Clemson, SC 29634, USA
 Email: <http://www.genome.clemson.edu/orders/index.html>
 This clone is available through the Clemson University Genomics Institute
 Seq primer: T3.

FEATURES

source

Location/Qualifiers

1..742
 /organism="Lycopersicon esculentum"
 /mol_type="mRNA"
 /cultivar="TA496"
 /db_xref="taxon:4081"
 /clone="CLB62H8"
 /cissue_type="Pericarp"
 /dev_stage="breaker"
 /lab_host="SOLR"
 /clone_1lb="tomato breaker fruit"
 /note="Vector: pBluescriptMCS2duplt; Site 1: EcoRI, Site 2: XhoI; supplier: Boyce Thompson Institute; sequencing: The Institute for Genomic Research. Fruit were harvested at the breaker stage (first sign of lycopen accumulation on the blossom end of fruit). Fruit were cut in half and the seeds and locules were discarded prior to freezing the pericarp."

ORIGIN

Alignment Scores:

Pred. No.: 4.6e-137 Length: 742
 Score: 1282.00 Matches: 246
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 53.64% Indels: 0
 DB: 4 Gaps: 0

US-10-691-374-2 (1-457) x BM413282 (1-742)

2 ValiieglnargAsenSerlleleuleuilellelePhAlaseSerlleSerThr 21
 3 GTTATCCAAAGGAAATGATTTCTCTTCATATTATTGCTTCATCAATTCACACT 62
 22 CybArGserAenValilleaPhaPhaenleuPhelyglnValiTyraPhaenilleu 41
 63 TGTAGAGCAATGATTGATGACAATTATTCAAACAGTTTATGATTAATCTTGAA 122
 42 GlnGluPhaAlaPhaPhaPhaPhaPhaPhaPhaPhaPhaPhaPhaPhaPha 61
 123 CAAGAAATTTGTCATGATTTTCACGTTATCTTTCTATTTCAGCAAAATTAATGAAAGC 182
 62 AenAenAenilleaPhlyAenlyAenlyAenlyAenlyAenlyAenlyAenlyAenly 81
 183 AAGCAATATTGACAAAGGTTGATTAATAATGGAATTAAGTATTAATGACTTACCTTT 242
 82 GlnAlaLygLyAenPGlyAenThTyraPhaenilleaPhaGlnGlnAlaTriPhaen 101
 243 GGAAGCTAAGGGTGAAGAAACATATGATTAATGATTCATTTGAGCAAGCATGATGA 302
 102 AlaCySerSerSerAenThProValGlnPhaValValProValAenlyAenlyAenly 121
 303 GATTTTCATCTAGAACACCTGTTCAATTTGGGTTCTTAAACAGAAATTAATCTTCTC 362
 122 LySGlnilleThPhaSerGlyProCybArGserSerlleSerVallyllePhaGlySer 141
 363 AAGCAATTCACCTTTTCAGGTCATGACATCTTCTATTTCAGTAAAGATTTTGGATCC 422
 142 LeuGlnAlaSerSerlylleSerAenlyTyraPhaPhaPhaPhaPhaPhaPhaPha 161
 423 TTAGAGAGCATCTAGTAATAATTCAGACCTCAAGATAGAAAGCTTTGATCTTTGAT 482
 162 SerValGlnAenleuValValGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 181
 483 AGTGTTCAAATTTGATGTTGAGAGGAGGAGCAACTCAAGGCAATGAGCAAGATGAG 542
 182 TTPProSerSerCybLySilleAenlyAenlyAenlyAenlyAenlyAenlyAenlyAenly 201

DB 543 TGGCCAAATTCTTGCAAAATAATAATCACTGCCATGACGAGATGACCAACAGGCTTA 602
 QY 202 ThrPhetPaaenCybLyAenleuLyValAenAenleuLySerLyAenAlaGln 221
 DB 603 ACCTTCGCAATTCGCAAAATTTGAAAGTGAATATCTTAAGAGATTAATGCAACACA 662
 QY 222 IlleHilleLyPhaGlnSerCybThraPhaenValAlaAseAenleuMerilleAenAla 241
 DB 663 ATTCATATCAAAATTTGAGTCATGACCTAATGTTGATCTTCAAAATTTGATGATCAATGCT 722
 QY 242 SerAlaLySerProAen 247
 DB 723 TCAGCAAGAGGCCAAAT 740

RESULT 6
 BM413343
 LOCUS
 DEFINITION EST587670 tomato breaker fruit Lycopersicon esculentum cDNA clone
 CLB63C11 5' end, mRNA sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Lycopersicon esculentum (tomato)
 Lycopersicon esculentum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 asterids; Lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.
 1 (bases 1 to 749)
 Alcalá, J., Vrebalov, J., White, R., Vialon, T., Karanymcheva, S.A.,
 Tsai, J., Bougri, O., Kirkness, E., Utecherback, T., Van Aken, S.,
 Roining, C.M., Fraser, C.M., Martin, G.B., Tanksley, S.D. and
 Giovannoni, J.

REFERENCE

AUTHORS

TITLE
 JOURNAL
 COMMENT
 Generation of ESTs from tomato fruit tissue, breaker stage (2002)
 Unpublished (2002)
 Contact: CUGI
 Clemson University Genomics Institute
 100 Jordan Hall, Clemson, SC 29634, USA
 Email: <http://www.genome.clemson.edu/orders/index.html>
 This clone is available through the Clemson University Genomics Institute
 Seq primer: T3.

FEATURES

source

Location/Qualifiers

1..749
 /organism="Lycopersicon esculentum"
 /mol_type="mRNA"
 /cultivar="TA496"
 /db_xref="taxon:4081"
 /clone="CLB63C11"
 /cissue_type="Pericarp"
 /dev_stage="breaker"
 /lab_host="SOLR"
 /clone_1lb="tomato breaker fruit"
 /note="Vector: pBluescriptMCS2duplt; Site 1: EcoRI, Site 2: XhoI; supplier: Boyce Thompson Institute; sequencing: The Institute for Genomic Research. Fruit were harvested at the breaker stage (first sign of lycopen accumulation on the blossom end of fruit). Fruit were cut in half and the seeds and locules were discarded prior to freezing the pericarp."

ORIGIN

Alignment Scores:

Pred. No.: 1.36e-136 Length: 749
 Score: 1278.00 Matches: 248
 Percent Similarity: 99.60% Conservative: 0
 Best Local Similarity: 99.60% Mismatches: 1
 Query Match: 53.47% Indels: 1
 DB: 4 Gaps: 0

US-10-691-374-2 (1-457) x BM413343 (1-749)

QY 74 LySValilleAenValleuSerPhaGlyAlaLygLyAenPGlyLyAenThTyraPhaenille 93

```

Db      3 AAGGTATTTATGTAATCTTACCTTGGAGCTTAAGGGTGAAGGAAAAAATATGATATAT 62
Qy      94 ALPhgGluGlnAlaTrpAsnGlnAlaCysSerSerArgThrProValGlnHeVal 113
Db      63 GCATTTGAGCAAGCATGGAATGAAGCATGTTCACTGAAACACCTGTTCAATTTTGCT 122
Qy      114 ProValAsnLysAsnLysLeuLeuLysGlnIleThrPheSerGlyProCysArgSer 133
Db      123 CCAAAAACAGAGATATCTTCTCAAGCAAAACACCTTTTCAGTCTCAATGAGATCTCT 182
Qy      134 ILeSerValLysIlePheGlySerLeuGlnAlaSerSerLysIleSerAspThr 153
Db      183 ATTTGAGTAAGATTTTGGATCTTGAAGAGCATTAAGTAATTTCAAGCTCAAGAT 242
Qy      154 ATGAGTLeuTrpIleAlaPheAspSerValGlnAsnLeuValGlyGlyGlyThr 173
Db      243 AAGAGGCTTTGAGATGCTTTGATAGTGTCAAAATTAGTTGAGGAGGAGAACT 302
Qy      174 ILeAsnGlyAsnGlyGlnValTrpTrpProSerSerCysLysIleAsnLysSerLeuPro 193
Db      303 ATCAATGGCAATGACCAAGTATGCGGCCAAGTCTTGCAAAATTAATTAATCACTGCCA 362
Qy      194 CysArgAspAlaProThrAlaLeuThrPheTrpAsnCysLysAsnLeuLysValAsnAsn 213
Db      363 TGCAAGGAGTGCACCAACGGCTTAACTTGAATTCGAAATTTGAAAGTGAATAT 422
Qy      214 LeuLysSerLysAsnAlaGlnGlnIleHisIleLysPheGlySerCysThrAsnVal 233
Db      423 CTAAAGGTAAATAATGCACAAATCAATATCAATTAATTAAGTCAAGCTAATGTTGTA 482
Qy      234 ALSerSerLeuMetIleAsnAlaSerAlaLysSerProAsnThrAspGlyValHisVal 253
Db      483 GCTTCAAAATTTGAGATCAATGCTTTCAGCAAGAGGCCAAATCTGATGAGTCCATGTA 542
Qy      254 SerAsnThrGlnTrpIleGlnIleSerAspThrIleIleGlyThrGlyAspAspCysIle 273
Db      543 TCAAAATCTCAATATATTCAAATATATGATATCAATTAATTAAGTCAAGCTAATGTTGTA 602
Qy      274 SerIleValSerGlySerGlnAsnValGlnAlaThrAsnIleThrCysGlyProGlyHis 293
Db      603 TCAATGCTTCTGATCTCAAAATGTCAGAGGCCCAAAATTTACTTGTTGCTCAAGTCTAT 662
Qy      294 GlyIleSerIleGlySerLeuGlySerGlyAsnSerGlnAlaTrpValSerAsnValThr 313
Db      663 GGTATAGATTTGAGAGCTTGAATCTGGAATTC-GAAGCTTATGTTCTAATGTTACT 721
Qy      314 ValAsnGlnAlaLysIleIleGlyAla 322
Db      722 GTAAATGAAGCCAAATTAATCGGTGCC 748

RESULT 7
AM441939 709 bp mRNA linear EST 18-MAY-2001
LOCUS EST11335 tomato fruit red ripe, TAWU Lycopersicon esculentum cDNA
DEFINITION
ACCESSION AM441939
VERSION AM441939.1 GI:6977190
KEYWORDS EST.
SOURCE Lycopersicon esculentum (tomato)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 709)
Alcala,J., Vrebalov,J., White,R., Matern,A.L., Holt,I.E., Liang,F.,
Upson,J., Hansen,T., Craven,M.B., Bowman,C.L., Ahn,S.,
Romling,C.M., Fraser,C.M., Martin,G.B., Tankeley,S.D. and
Giovannoni,J.
Generation of ESTs from tomato fruit tissue
Unpublished (1999)
CONTACT: CUGI
Clemson University Genomics Institute

```

FEATURES

Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>
5 prime sequence.
Location/Qualifiers

1..709

/organism="Lycopersicon esculentum"

/mol_type="mRNA"

/cultivar="TA496"

/db_xref="taxon:4081"

/clone="CLEN19819"

/issue="pericarp"

/dev_stage="red ripe (7-20 days post-breaker)"

/clone_1lb="tomato fruit red ripe, TAWU"

/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; supplier: Giovannoni; Fruit were tagged at the breaker stage (first sign of lycopen accumulation on the blossom end of the fruit) and harvested 7 days post-breaker (fully red-ripe). 10 days post breaker, and 20 days post-breaker (over-ripe). 20 day fruit which showed external or internal signs of pathogenesis were discarded. Fruit were cut in half and the seeds and locules were discarded prior to freezing the pericarp."

ORIGIN

Alignment Scores:

Pred. No.: 6,31e-130 Length: 709
Score: 1220.00 Matches: 235
Percent Similarity: 99.58% Conservative: 0
Best Local Similarity: 99.58% Mismatches: 1
Query Match: 51.05% Indels: 0
DB: 2 Gaps: 0

US-10-691-374-2 (1-457) x AM441939 (1-709)

```

Qy      219 ALaGlnGlnIleHisIleLysPheGlySerCysThrAsnValAlaSerAsnLeuMet 238
Db      2 GCACAAACAAATTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 61
Qy      239 ILeAsnAlaSerAlaLysSerProAsnThrAspGlyValHisValSerAsnThrGlnTrp 258
Db      62 ATCAATGCTTCAAGCAAGGCCCAAAATCTGATGAGTCAAGTATCAATCTCAATTTGAG 61
Qy      259 ILeGlnIleSerAspThrIleIleGlyThrGlyAspAspCysIleSerIleValSerGly 278
Db      122 ATTCAAATATCTGATATCTATTTATGGAACAGGATGATGATTTCAATGTTCTGGA 181
Qy      279 SerGlnAsnValGlnAlaThrAsnIleThrCysGlyProGlyHisGlyIleSerIleGly 298
Db      182 TCTCAAAATGTCAGAGCCCAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 241
Qy      299 SerLeuGlySerGlyAsnSerGlnAlaTrpValSerAsnValThrValAsnGlnAlaLys 318
Db      242 AGCTTGAAGATCTGGAATTCAGAACTTATGTTGTTAATGTTAATGAAGCCAA 301
Qy      319 ILeIleGlyAlaGlyAsnGlyValAlaGlyIleLysThrTrpGlnGlyGlySerGlyGlnAla 338
Db      302 ATTATCGGTGCGCAAAATGAGTTAGATCAAGACTTGGCAGGAGATTCGACCAAGCT 361
Qy      339 SerAsnIleLysPheLeuAsnValGlnMetGlnAspValLysTrpProIleIleLysAsp 358
Db      362 ACACAATCAAAATTTCTGAATGTGGAATGCAAGACTTGAAGTATCCCAATTAATTAATGAC 421
Qy      359 GlnAsnTrpCysAspArgValGlnProCysIleGlnGlnPheSerAlaValGlnValLys 378
Db      422 CAAAATCTATGATCGATCGATGGAACCAATATCAACAAAGTTTTCAGCAGATTAAGTGA 461
Qy      379 AsnValValTrpGlnAsnIleLysGlyThrSerAlaThrLysValAlaIleLysPheAsp 398
Db      482 AATGTGTGTATGGAATATCAAGGCAAGAGTCAACAAAGGTGGCATTAATTAATTTGAT 541
Qy      399 CysSerThrAsnProCysGlyGlyIleIleMetGlnAsnIleAsnLeuValGlyGln 418

```

Db 542 TGACACAACTTCATGTGTAAGAAATTAATGAGAAATTAATTAGTAAGGAA 601

Qy 419 SerGlyLyProSerGlyLysAlaThrCysIleValAsnValHisPheAsnAsnAlaGluHisVal 438

Db 602 AGTGAACCAACATCAAGAGCTAGCTCAAAAAATGTCATTTTAAACATGCTGAACATGTT 661

Qy 439 ThrProHisCysThrSerLeuGluIleSerGlyLysAlaGluLysLeu 454

Db 662 AACACCACTGCACTTCACTAGAAATTTCAAGAGATGAAGCTCTTTTG 709

RESULT 8
AM442253 696 bp mRNA linear EST 18-MAY-2001
LOCUS EST311649 tomato fruit red ripe, TAMU Lycopersicon esculentum cDNA
DEFINITION clone cLEN22G14 5', mRNA sequence.
ACCESSION AM442253
VERSION AM442253.1 GI:6977504
KEYWORDS EST.
SOURCE Lycopersicon esculentum (tomato)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 696)
Alcala,J., Vrebalov,J., White,R., Matern,A.L., Holt,I.E., Liang,F.,
Upton,J., Hansen,T., Craven,M.B., Bowman,C.L., Ahn,S.,
Romling,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and
Giovannoni,J.
Generation of ESTs from tomato fruit tissue
Unpublished (1999)
CONTACT: CUGI
Clemson University Genomics Institute
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>
5 prime location.
FEATURES
source
1..696
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEN22G14"
/cissue_type="pericarp"
/dev_stage="red ripe (7-20 days post-breaker)"
/clone_lib="tomato fruit red ripe, TAMU"
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; supplier: Giovannoni; Fruit were tagged at the
breaker stage (first sign of lycopen accumulation on the
blossom end of the fruit) and harvested 7 days
post-breaker (fully red-ripe), 10 days post breaker, and
20 days post-breaker (over-ripe). 20 day fruit which
showed external or internal signs of pathogenesis were
discarded. Fruit were cut in half and the seeds and
locules were discarded prior to freezing the pericarp."

ORIGIN
Alignment Scores:
Pred. No.: 5.16e-129 Length: 696
Score: 1212.00 Matches: 231
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 50.71% Indels: 0
DB: 2 Gaps: 0

US-10-691-374-2 (1-457) x AM442253 (1-696)

Qy 101 GUAAGCysSerSerGlyThrProValGlnPheValProLysAsnLysAsnTyrLeu 120

Db 3 GAAGCAAGTTCATCTAGAACACCTGTTCAATTGGTTCATAAACAAGAAATATCTT 62

Qy 121 LeuLYGlnIleThrPheSerGlyProCysArgSerSerIleSerValLysIlePheGly 140

Db 63 CTCAGCAAAATCACCTTTTCAGGTCCATGCAGATCTTATTTCAAGTAAGATTTTGG 122

Qy 141 SerLeuGlnLysSerSerIleIleSerAspTyrLysAspArgLeuTyrIleAspLe 160

Db 123 TCCCTTAAGAAAGCATCTAGTAATAATTTTCAGCTCAAAAGATGAAGCTTTGGATTGCTTTT 182

Qy 161 AspSerValGlnAsnLeuValValGlyGlyGlyThrIleAsnGlyAsnGlyVal 180

Db 183 GATAGCTTCAAAATTTTACTTCTTGAGAGAGAGAGAACTATCATATGCAGATGGAACAAGTA 242

Qy 181 TyrTrpProSerSerCysLysIleAsnLysSerLeuProCysArgAspAlaProThrAla 200

Db 243 TCGTGCCCAAGTCTCTGCAAAATTAATCACTGCCATGCAAGGATCCACCACGGCC 302

Qy 201 LeuThrPheThrAspCysIleValAsnLysValAsnAsnLeuLysSerLysAsnAlaGln 220

Db 303 TTACCTCTTGAAATTCGAAAAATTTGAAAGTGAATTAATCTAAAGAGTAATAATGACAA 362

Qy 221 GlnIleHisIleLysPheGluSerCysThrAsnValValAlaSerAsnLysMetIleAsn 240

Db 363 CAATTCATATCAATTAATTTAGTCATGCACTAATGTTGATGCTTCAATTTGATGATCAT 422

Qy 241 AlaSerAlaLysSerProAsnThrAspGlyValHisValSerAsnThrGlnTyrIleGln 260

Db 423 GCTTCAGCAAAAGCCCAAAATCTGATGAGATGCCATGTATCAAAATTAATATATTCAA 482

Qy 261 IleSerAspThrIleIleGlyThrGlyAspAspCysIleSerIleValSerGlySerGln 280

Db 483 ATATCTGATACATTAATTTAGAAACAGTGATGATGATTTCAATGTTCTGATCTCAA 542

Qy 281 AsnValGlnAlaThrAsnIleThrCysGlyProGlyHisGlyIleSerIleGlySerLeu 300

Db 543 AATGTCAGAGCCCAAAATTTACTTGTGTCACAGATCAGATGATTAAGTTTGAAGCTTA 602

Qy 301 GlySerGlyAsnSerGlyLysAlaTyrValSerAsnValThrValAsnGlnLysIleIle 320

Db 603 GGATCTGGAATTTCAAGACCTATGCTATGTAATGTAATGAATGAACCAAAATTAATTC 662

Qy 321 GlyAlaGlnLysGlyValArgIleLysThrTyr 331

Db 663 GGATGCCGAAATGAGATTAGATCAAGACTTG 695

RESULT 9
AM442335 687 bp mRNA linear EST 18-MAY-2001
LOCUS EST311731 tomato fruit red ripe, TAMU Lycopersicon esculentum cDNA
DEFINITION clone cLEN22L13 5', mRNA sequence.
ACCESSION AM442335
VERSION AM442335.1 GI:6977586
KEYWORDS EST.
SOURCE Lycopersicon esculentum (tomato)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 687)
Alcala,J., Vrebalov,J., White,R., Matern,A.L., Holt,I.E., Liang,F.,
Upton,J., Hansen,T., Craven,M.B., Bowman,C.L., Ahn,S.,
Romling,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and
Giovannoni,J.
Generation of ESTs from tomato fruit tissue
Unpublished (1999)
CONTACT: CUGI
Clemson University Genomics Institute
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>
5 prime location.
FEATURES
source
1..687
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"

/db_xref="taxon:4081"
/clone="clEN22C13"
/tissue_type="pericarp"
/dev_stage="red ripe (7-20 days post-breaker)"
/clone_lib="tomato fruit red ripe, TAMU"
/note="Vector: pBluescript SK(-); Site_1: EcoRI, Site_2:
XhoI; supplier: Giovannoni, Fruit were tagged at the
breaker stage (first sign of lycopene accumulation on the
blossom end of the fruit) and harvested 7 days
post-breaker (fully red-ripe), 10 days post breaker, and
20 days post-breaker (over-ripe). 20 day fruit which
showed external or internal signs of pathogenesis were
discarded. Fruit were cut in half and the seeds and
locules were discarded prior to freezing the pericarp."

ORIGIN

Alignment Scores:

Pred. No.:	1,046-126	Length:	687
Score:	1192.00	Matches:	226
Percent Similarity:	99.56%	Conservative:	1
Best Local Similarity:	99.12%	Mismatches:	1
Query Match:	49.87%	Indels:	0
DB:	2	Gaps:	0

US-10-691-374-2 (1-457) x AM442335 (1-687)

QY 76 llaanvalleuSerPheGlyAlaGlySerGlyThrTyraAspAnlleAlaPhe 95
DB 3 ATTATGTACTTAACTTTGAAAGCTTAGGCTATGGAAGAAATGAAATATTCATTT 62
QY 96 GluGlnAlaTTPaenGluAlaCySerSerArgThrProValGlnPheValProLys 115
DB 63 GAGCAAGATGAAATGAAAGCATGTTCATAGAAACACCTTCATTGTTGTTCCAAA 122
QY 116 AasnlySantyrleuLeuLysGlnlleThrPheSerGlyProCysArgSerlleSer 135
DB 123 AACAAAGATATCTTCCAGCAAAATCACTTTACAGTTCACAGCAATCTTATTTCA 182
QY 136 VallySlePheGlySerleuGluAlaSerSerlySleSerAspTyrllySAspArg 155
DB 183 GTTAAAGATTTTGGATCTTAGAAGCATCTAGAAATTTCACTACAAAGATGAAAG 242
QY 156 leuTrpIleAlaPheAspSerValGlnAsnLeuValGlyGlyGlyThrIleAsn 175
DB 243 CTGTGATGCTTTGATGATGTTCAAAATTTTGGTGGAGAGAGAGAACTATCAT 302
QY 176 GlyAsnGlyGlnValTrpTrpProSerSerCyslySleAsnLysSerleuProCysArg 195
DB 303 GGCATATGACAAATGATGTGGCCAAAGTCTTGCAAAATTAATTAATCACTGCATGCAG 362
QY 196 AspAlaProThrAlaLeuThrPheTrpAsnCylySAsnLeuLysValAsnAsnLys 215
DB 363 GATGACCAACGGCTTAACTTCTGAAATTCAAAATTTGAAAGTCAATATCTTAAG 422
QY 216 SerLySAsnAlaGlnGlnlleHslylePheGlySerCysThAsnValValAlaSer 235
DB 423 AGTAAAGAAAGCAACAATTCATATCAATTTGATGATGATCACTAACTTGTAGCTTCA 482
QY 236 AasnLeuMetlleAsnAlaSerAlaLysSerProAsnTrpAspGlyValHsValSerAsn 255
DB 483 AATTGATGATCAATGCTTCAAGCAAGAGCCCAATTAATCTGATGGAGTCAATGATCAAT 542
QY 256 ThrGlnTyrlleGlnlleSerAspThrIlelleGlyThrGlyAspAspCyslleSerlle 275
DB 543 ACTCAATATATCAATATCTGATCTATTTATGGAACAGGTGATGATGATTTCAAT 602
QY 276 ValSerGlySerGlnAsnValGlnAlaThrAsnIleThrCysGlyProGlyHsGlylle 295
DB 603 GTTTCTGATCTCAAAATGTGACAGCCACACATATTTCTTGTGTCCAGGTCAATGATATA 662
QY 296 SerlleGlySerleuGlySerGly 303
DB 663 AGTATGGAAGCTTAGATCTGGA 686

RESULT 10

AM442241

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITILE

JOURNAL

COMMENT

FEATURES

source

1..677

/organism="Lycopersicon esculentum"

/mol_type="mRNA"

/cultivar="74496"

/db_xref="taxon:4081"

/clone="clEN22C16"

/tissue_type="pericarp"

/dev_stage="red ripe (7-20 days post-breaker)"

/clone_lib="tomato fruit red ripe, TAMU"

/note="Vector: pBluescript SK(-); Site_1: EcoRI, Site_2:
XhoI; supplier: Giovannoni, Fruit were tagged at the
breaker stage (first sign of lycopene accumulation on the
blossom end of the fruit) and harvested 7 days
post-breaker (fully red-ripe), 10 days post breaker, and
20 days post-breaker (over-ripe). 20 day fruit which
showed external or internal signs of pathogenesis were
discarded. Fruit were cut in half and the seeds and
locules were discarded prior to freezing the pericarp."

US-10-691-374-2 (1-457) x AM442241 (1-677)

QY 189 AasnlySerleuProCysArgAspAlaProThrAlaLeuThrPheTrpAsnCylySAsn 208

DB 1 AATTAATCTGCTGCAATGAGGATGCAACACGGCTTAACTTCTGGAATTCGAAAAT 60

QY 209 LeuLysValAsnAsnLeuLysSerlySAsnAlaGlnGlnlleHslylePheGlySer 228

DB 61 TTGAAGAGGATTAATCTTAAGAGTAAATTCACAAATTCATATCAATTTGAGATCA 120

QY 229 CysThrAsnValValAlaSerAsnLeuMetlleAsnAlaSerAlaLysSerProAsnThr 248

DB 121 TGCATTAATGTTGAGCTTCAAAATTTGATGATCAATCTTCAAGAAAGACCCAAATCT 180

QY 249 AspGlyValHsValSerAsnThrGlnTyrlleGlnlleSerAspThrIlelleGlyThr 268

DB 181 GATGAGTCCATGATCAATTAATCTCAATATTCATATCTGATATCTATTATGGAACA 240

QY 269 GYAAPAPGCGYGLIeser11eValSerGlySerGlnAsnValGlnAlaThrAsn1Iethr 288
DB 241 GGATGAGATTGATTTCAATTGTTCTGGATCTCAAAAGTGCAGGCCCAATATTTACT 300
QY 289 CYSGLYPRGGLYH1eG1YIleSer1IeG1YSerLeuG1YSerG1YAsnSerGlnAlaTYr 308
DB 301 TGTGCTCCAGCTCATGCTAATGATTTGGAGCTTGAGATTCGGAATTCAGAGCTTAT 360
QY 309 ValSerAsnValThrValAsnGlnAlaValYIle1IeG1YAlaGlnAsnG1YValArg1Ie 328
DB 361 GGTCTAATAGTTACTGTAATATGACCAAAATATATCGTGCCGAAATATGAGTTAGATC 420
QY 329 YserThrThrGlnG1YG1YSerG1YGlnAlaSerAsn1IeValPheLeuAsnValGlnUmc 348
DB 421 AAGACTGGCAGGAGGAGATCTGCACAGCTAGCAACATCAAAATTTCTGAAATGCGAAAG 480
QY 349 GlnAepValYserTYrPro1Ie1Ie1IeAepGlnAsnTYrCYeAPArgValGlnProCYs 368
DB 481 CAAGAGCTTAAGTATCCATTAATTAAGACCAAAATATTTGATGATGATGACCATCT 540
QY 369 1IeGlnGlnPheSerAlaValGlnValYAsnValValTYrGlnUbn1IeYsg1YThr 388
DB 541 ATACAAACAGTTTTCACAGCTTCAAGTGAATAATGTGTATGAGATATCAAGGCGACA 600
QY 389 SerAlaThrValValAla1IeYsPheAPCYeSerThrAsnPheProCYsG1YUc1YIle 408
DB 601 AGTGCACAAAGTGGCCATTAATTTGATGACGACAAACCTTTCATGTGAAAGATT 660
QY 409 1IeMerGlnUbn1Ie 413
DB 661 ATATGAGATTA 675

RESULT 11

BM408546 675 bp mRNA linear EST 22-JAN-2002
LOCUS EST82873 tomato breaker fruit Lycopersicon esculentum cDNA clone
DEFINITION CLE645C7 5' end, mRNA sequence.

ACCESSION BM408546
VERSION BM408546.1 GI:18260176
KEYWORDS EST.

SOURCE Lycopersicon esculentum (tomato)
ORGANISM Lycopersicon esculentum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 675)

REFERENCE

Ahalla, J., Vrebalov, J., White, R., Vision, T., Karimicheva, S.A.,
Roming, C.M., Frazer, C.M., Martin, G.B., Tanksley, S.D. and
Giovannoni, J.
Generation of ESTs from tomato fruit tissue, breaker stage (2002)
Unpublished (2002)
Contact: CUGI

TITLE

Clemson University Genomics Institute

100 Jordan Hall, Clemson, SC 29634, USA

Email: <http://www.genome.clemson.edu/orders/index.html>

This clone is available through the Clemson University Genomics

Institute

Seq primer: T3.

FEATURES

Source

Location/Qualifiers
1..675
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="CLE645C7"
/issue_type="Pericarp"
/dev_stage="breaker"
/lab_host="SOLR"
/clone_lib="tomato breaker fruit"
/note="Vector: pBluescriptskmCudapc; site_1: EcoRI;

Site 2: XhoI; supplier: Boyce Thompson Institute;
sequencing: The Institute for Genomic Research. Fruit
were harvested at the breaker stage (first sign of
lycopene accumulation on the blossom end of fruit). Fruit
were cut in half and the seeds and locules were discarded
prior to freezing the pericarp."

ORIGIN

Alignment Scores:

	Pred. No.:	Length:	Score:
	5,06e-123	675	1160.00
Score:	99.11%	Matches:	222
Percent Similarity:	99.11%	Conservative:	0
Best Local Similarity:	99.11%	Mismatches:	2
Query Match:	48.54%	Indels:	0
DB:	4	Gaps:	0

US-10-691-374-2 (1-457) x BM408546 (1-675)

QY 104 SerSerArgThrProValGlnPheVal1ProLYAsnLYAsnTYrLeuLYsgln 123
DB 2 TCATCTAGAACCTGTTCAATTTGTGTTCTTAACCAAAATTAATTTCTCAAGCA 61
QY 124 1IethrPheSerG1YProCYsArgSer1IeSerValYIlePheG1YSerLeuGln 143
DB 62 ATCACCCTTTTTCAGTCCATGACGATCTTCTATTCAGTAAAGATTTTGGATCCTTAGAA 121
QY 144 1IaSerSerLYe1IleSerAPTYrLYsAPArg1IeUtr1IeAlaPheAPSerVal 163
DB 122 GCATCTAGAAATTTTCAGCTACCAAGATACAGCTTGGATTCCTTTATATGTT 181
QY 164 GlnAenLeuValValG1YG1YLYr1IeAsnG1YQlnVal1TrpTrpPro 183
DB 182 CAAATTTAGTTGTTGAGAGAGAGAACTTCAATGCAATGACAAATATGTTGCCA 241
QY 184 SerSerCYeLYe1IeAsnLYSerLeuProCYsArgAPAlaProThrAlaLeuThrPhe 203
DB 242 ACTTCTGCAAAATTAATTAATCACTGCCATGACGAGGACCAACGCTTAACCTTC 301
QY 204 TrpAsnCYeLYeAsnLeuLYsValAsnLeuLYsSerLYsAsn1IeGln1IeHis 223
DB 302 TCGAATTCGCAAAATTTGAAAGTGAATTAATCAAGATTAAGTCAACAAATTCAT 361
QY 224 1IeLYsPheGlnUbnCYeThrAsnValVal1IeSerAsnLeuMer1IeAsn1IeSerAla 243
DB 362 ATCAATTTAGATCACTGATATGTTAGCTTCAATTTATGATGATCAATGCTTAGCA 421
QY 244 LYSerProAsnThrAspG1YVal1IeVal1IeSerAsnThrGlnTYr1IeGln1IeSerAP 263
DB 422 AAGACCCCAATACATGATGAGATCCATGATCAATATCAATATATCTGAT 481
QY 264 Thr1Ie1IeG1YThrG1YAspAPCYe1IeSer1IeVal1IeSerG1YSerGlnAsnValGln 283
DB 482 ACTATTTATGGAACAGTATGATTTCAATTTCTCGATCTCAAAATGCGCAG 541
QY 284 1IaThrAsn1IeThrCYeG1YProG1YH1eG1Y1IeSer1IeG1YSerLeuG1YSerG1Y 303
DB 542 GCCACCAATATTAATCTGCTGCTCAGGTCATGATTAAGTATTAAGTATGAGATTCGA 601
QY 304 AsnSerGlnAlaTYrValSerAsnValThrValAsnGlnAlaLYe1Ie1IeG1YAlaGln 323
DB 602 AATTCAGAGCTTAATGTTCTAATCTTACTGTAATGAAGCAAAATTAATCGGTGCCGA 661
QY 324 AsnG1YValArg 327
DB 662 AATGAGATTAGG 673

RESULT 12
AW221785 672 bp mRNA linear EST 18-MAY-2001
LOCUS EST98596 tomato fruit red ripe, TAMU Lycopersicon esculentum cDNA
DEFINITION clone cLEM3122, mRNA sequence.
ACCESSION AW221785
VERSION AW221785.1 GI:6533469

KEYWORDS EST.
SOURCE Lycopersicon esculentum (tomato)
ORGANISM Lycopersicon esculentum
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon. 1 (bases 1 to 672)
AUTHORS Alcala,J., Vrebalov,J., White,R., Matern,A.L., Holt,I.E., Liang,F., Upton,J., Hansen,T., Craven,M.B., Bowman,C.L., Ahn,S., Rønning,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.
TITLE Generation of ESTs from tomato fruit tissue
JOURNAL Unpublished (1999)
COMMENT Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>
5 prime sequence.

FEATURES
source
1..672
Location/Qualifiers
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="CLEN3122"
/tissue_type="pericarp"
/dev_stage="red ripe (7-20 days post-breaker)"
/clone_lib="tomato fruit red ripe, TAMU"
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; supplier: Giovannoni; Fruit were tagged at the blossom end of the fruit) and harvested 7 days post-breaker (fully red-ripe). 10 days post breaker, and 20 days post-breaker (over-ripe). 20 day fruit which showed external or internal signs of pathogenesis were discarded. Fruit were cut in half and the seeds and locules were discarded prior to freezing the pericarp."

ORIGIN

Alignment Scores:
Pred. No.: 9.77e-123 Length: 672
Score: 1157.50 Matches: 221
Percent Similarity: 99.10% Conservative: 0
Best Local Similarity: 99.10% Mismatches: 1
Query Match: 48.43% Indels: 1
DB: 2 Gaps: 1

US-10-691-374-2 (1-457) x AW221785 (1-672)

QY 90 TTTATGATTAATTTGCAATTTGAGCAAGATGGAATGAAAGCATGTTCACTAGAACACCTGTT 109
DB 3 TATGATTAATTTGCAATTTGAGCAAGATGGAATGAAAGCATGTTCACTAGAACACCTGTT 62
QY 110 GlnpheValValProLysAsnLysAsnTYLeuLeuLysGlnLLeuPheSerGlyPro 129
DB 63 CAATTTGTTGTTCTTAATAAAGATTTCTTCTCAAGCAATCACTTTTCAGGTCA 122
QY 130 CysArgSerSerLysSerValLysLysLeuPheGlySerLeuGlnLLeuLysSerLysLysSer 149
DB 123 TGGAGATCTTCTTAATTCAGTAAAGATTTTGGATCTTCAAGCAATCACTTTTCAGGTCA 182
QY 150 AspTYLysAspArgLysLeuTrpLysLeuLysPheAspSerValGlnAsnLeuValGly 169
DB 183 GACTACAAAGATAGAAGCTTTGATGCTTTGATAGTTCATAAATTTAGTTGGA 242
QY 170 GlyGlyGlyThrLLeuGlnGlyValGlnValTrpProSerSerCysLysLLeu 189
DB 243 GAGGAGGAGTATCAATGAGCAATGACAAAGATGAGTGGCAAGTTCTTGCAAAATTAAT 302
QY 190 LysSerLeuProCysArgAspAlaProThrAlaLeuThrPheTrpAsnGlyAsnLeu 209
DB 303 AAATCATGCGCATGAGGAGATGACCAAGCGCTTAACCTTACCTTGTGAATTCAGAAATTTG 362

QY 210 LysValAsnAsnLeuLysSerLysAsnAlaGlnGlnLLeuLysPheGlySerCys 229
DB 363 AAAGTGAATATCTAAAGATGAAAATGCAACAATTCATATCAATTTAGTCATCC 422
QY 230 ThrAsnValValLysSerLeuMetLLeuAlaLLeuLysSerProAsnThr 248
DB 423 ACTAATCTTGTACTTCAATTTGATGATCAATGCTTTTTCAGCAAGAGCCCAATTAAT 482
QY 249 AspGlyValHisValSerAsnThrGlnTYrLLeuGlnLLeuSerAspThrLLeuGlyThr 268
DB 483 GATGAGTCCATGATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 542
QY 269 GlyAspAspCysLysSerLysValSerGlySerGlnAsnValGlnLLeuLysSerLys 288
DB 543 GGTGATGATTTGATTTCAATTTGATGATCAATGCTTTTTCAGCAAGAGCCCAATTAAT 602
QY 289 CysGlyProGlyHisGlyLLeuSerLLeuGlySerLeuGlyAsnSerGlyLysSerGly 308
DB 603 TGTGTCAGTCATGATATAGTATTTGGAAGCTTGAATCTGAAATTCAGAAAGCTTAT 662
QY 309 ValSerAsn 311
DB 663 GTGTCTAAT 671

RESULT 13
AW223400 643 bp mRNA linear EST 18-MAY-2001
LOCUS EST300211 tomato fruit red ripe, TAMU Lycopersicon esculentum cDNA
DEFINITION clone CLEN11114, mRNA sequence.
ACCESSION AW223400
VERSION AW223400.1 GI:535084
KEYWORDS EST.
SOURCE Lycopersicon esculentum (tomato)
ORGANISM Lycopersicon esculentum
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon. 1 (bases 1 to 643)
AUTHORS Alcala,J., Vrebalov,J., White,R., Matern,A.L., Holt,I.E., Liang,F., Upton,J., Hansen,T., Craven,M.B., Bowman,C.L., Ahn,S., Rønning,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.
TITLE Generation of ESTs from tomato fruit tissue
JOURNAL Unpublished (1999)
COMMENT Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>
5 prime sequence.

FEATURES
source

1..643
Location/Qualifiers
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="CLEN11114"
/tissue_type="pericarp"
/dev_stage="red ripe (7-20 days post-breaker)"
/clone_lib="tomato fruit red ripe, TAMU"
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; supplier: Giovannoni; Fruit were tagged at the blossom end of the fruit) and harvested 7 days post-breaker (fully red-ripe). 10 days post breaker, and 20 days post-breaker (over-ripe). 20 day fruit which showed external or internal signs of pathogenesis were discarded. Fruit were cut in half and the seeds and locules were discarded prior to freezing the pericarp."

ORIGIN

Alignment Scores:

Pred. No.:	2.35e-119	length:	644
Score:	1120.00	Matches:	213
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	47.20%	Indels:	0
DB:	2	Gaps:	0

US-10-691-374-2 (1-457) X AW223400 (1-643)

Oy	90	Pyrrapamni1lea1Phegiugi1na1a1Tpaangua1a1Cyseaser1qr1TProva1	109
Db	3	TATGTAATATTGCTCATTTTACAGACATGAAATGAAGCATGTTCACTTGAACACCTGTT	62
Oy	110	GlnPheVal1a1ProLyAsn1LyAsn1r1Leu1Leu1yEg1n1e1ThrPheSer1y1Pro	129
Db	63	CAATTGTGGTTCCTTAATAACAAGATTATCTTCTCAACCAATACCTTTCAAGTCCA	122
Oy	130	Cy0A1gSerSer11aSerVal1y1ePhegi1ySer1eug1ua1aSer1e1y11eSer	149
Db	123	TGCAGATCTTCTATTTCAGTAAAGATTTTGGATCCTTGAAGACATCTGTAAATTTCA	182
Oy	150	Aap1r1y1yAspA1yA1g1e1u1T1p11ea1aPheAapSer1Val1Glna1e1u1Val1G1y	169
Db	183	GACTACAAAGATGAAGCGCTTGGATGCTTTTGAATGTTCCAAATTTAGTGTGGA	242
Oy	170	G1y1g1yG1y1Th11a1a1nG1yA1nG1yG1n1a11T1p1T1P1roSerSer1Cy1y11e1a1n	189
Db	243	GAGAGAGAACTATCATATGCAATGGCAATGACAAAGTATGTTGGCCAAATCTTGCAAAATTAAT	302
Oy	190	Ly0Ser1e1u1Pro1CyA1rG1AspA1a1Pro1Th1a1a1e1u1Th1Phe1Tpa1n1Cy1y1a1e1u1	209
Db	303	AAATCAGCTGCCATGACGAGATGCCACCAACGGCCTTAACTTCTGAAATGGCAAAATTTTG	362
Oy	210	Ly0Val1a1a1n1a1e1u1y1Ser1y1a1n1a1a1nG1n11e1H1e1r1e1yA1p1hegi1u1Ser1Cy1s	229
Db	363	AAAGGAATTAATCTTAAGAGTAAATGAACAAACAAATTCATATCAATTTGAGTCATGC	422
Oy	230	Thra1n1Val1a1a1Ser1a1n1e1u1e1c11e1a1n1a1a1Ser1Ala1y1Ser1Pro1a1n1Th1rA1p	249
Db	423	ACTAATGTTGATGCTTCAATTTGATGATCAATGCTTCGCAAAAGGCCAAATATCTGAT	482
Oy	250	G1yVal1H1a1e1a1Se1a1n1Th1rG1n1r11eG1n11e1Se1A1p1Th1r11e1r11eG1y1Th1rG1y	269
Db	483	GGAGCCCATGTTCAATATCTCAATATATCTCAATATCTGATACATTTATTGGACAGGT	542
Oy	270	A1pA1p1Cy11e1Se1r11e1Val1e1rG1y1Se1rG1n1a1n1Val1Gln1a1Th1rA1n11e1Th1rCy1s	289
Db	543	GATGATTTGATTTCAATGTTCTTCTGGATCTCAAAAGTGCACAGGCCCAATATTTACTTGT	602
Oy	290	G1y1P1roG1yH1a1G1y11e1Se1r11eG1y1Se1r1eug1y1Ser	302
Db	603	GGTCCAGATCATGATTAAGTATGTAAGCTTAAGATCT	641
RESULT 14			
LOCUS	AM222222	672 bp	mRNA EST 18-MAY-200
DEFINITION	EST299033 tomato fruit red ripe, TAMU Lycopersicon esculentum cDNA clone c1EWN716, mRNA sequence.		
ACCESSION	AM222222		
VERSION	AM222222.1	GI:6533906	
KEYWORDS	EST.		
SOURCE	Lycopersicon esculentum (tomato)		
ORGANISM	Lycopersicon esculentum		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; Lamiales; Solanales; Solanaceae; Solanum; Lycopersicon. 1 (bases 1 to 672)		
AUTHORS	Alcala,J., Vrebalov,J., White,R., Matern,A.L., Hclt,I.E., Liang,F., Upcon,J., Hansen,T., Craven,M.B., Bowman,C.L., Ahn,S., Roming,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.		
TITLE	Generation of ESTs from tomato fruit cDNA		
JOURNAL	Unpublished (1999)		

COMMENT

Contact: **CGI**
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>
5 prime sequence. *Genetics*

FEATURES

Source

```

/organism="Lycopersicon esculentum"
/mol_type="mrna"
/cultivar="RA496"
/db_xref="taxon:4081"
/clone="CLEN716"
/tissue_type="pericarp"
/dev_stage="red ripe (7-20 days post-breaker)"
/clone_lib="tomato fruit red ripe, TAMU"
/notes="vector: pInuScript SK(-); Site_1: EcoRI, Site_2:
XhoI; supplier: Giovannoni; Fruit were tagged at the
breaker stage (first sign of lycopersin accumulation on the
blossom end of the fruit) and harvested 7 days
post-breaker (fully red-ripe), 10 days post breaker, and
20 days post-breaker (over-ripe). 20 day fruit which
showed external or internal signs of pathogenesis were
discarded. Fruit were cut in half and the seeds and
locules were discarded prior to freezing the pericarp."

```

ORIGIN

Alignment Scores:

Pred. No.:	3.6e-118	Length:	672
Score:	118.00	Matches:	216
Percent Similarity:	99.54%	Conservative:	0
Best Local Similarity:	99.54%	Mismatches:	0
Query Match:	46.78%	Indels:	1
DB:	2	Gaps:	0

US-10-691-374-2 (1-457) x AW222222 (1-672)

Oy	1	MeValllleglmaRphasnSerllleuLeuLeuIlellelePhalSesSerllSer	20
Dd	24	ATGGTATTCAGAAAGAAAGATGATTCCTTCCTCATATTATTTGGCTTCATCAATTC	83
Oy	21	ThrCysRgSerxanVallleaPaPaPaenLeuPheYsgInValTyrAspAsnIleLeu	40
Dd	84	ACTGTGGAAGCAATGTTATGTATGATGACATTTATTCAAAACAAGTTATGATTAATTCCTT	14
Oy	41	GlUGlGlnPhehIaIaPaPheGlnaIaTyrLeuSerTyrLeuSerTyrAsnIleGlu	60
Dd	144	GAACAAAGAAATGGCTCATATATTTCAAGCTTATCTTTATTATGACAAAAATATGA	20
Oy	61	SerAsnAsnAsnIleAspIyValAspIyAsnGlylleYsValleAsnValLeuSer	80
Dd	204	AGCAACAATATATTGACAAAGGTGATAAAAATGGATTAAGATGATTAACTACTTAC	26
Oy	81	PheGlyValAlaYsgIyAspGlyIySerThrTyrAspAsnIleAlaPheGluGlnAlaTyrAsn	100
Dd	264	TTTGGAGCTTAAGGCTGATGAAAAACATATGTATATATTGCAATTTGAGCAAGCAATG	32
Oy	101	GlAlaAcYsSerSerxThrProValGlnPheAlaValProYsAsnIyAsnTyrLeu	120
Dd	324	GAAGCATGTTCACTGAGAACCTGTTCAATTTGGTTCCATAAACAAGATTATCTT	38
Oy	121	LeuIyGlnIleThrPheSerGlyProCYsRgSerSerIleSerValIyGllePheGly	140
Dd	384	CTCAAGCAATCACTTTTCAGGTCCATGCAATCTTCTATTTCAAGTAAAGATTTTGTGA	44
Oy	141	SerLeuGlnIaIaSerSerIyIleSerAspTyrIyAspAsnArgGluTyrIleAlaPhe	160
Dd	444	TCCTTAGAAGCACTGAGTAATTTCAAGCTCAAAAGATAGAAAGCTTTGGATTCCTTT	50
Oy	161	AspSerValGlnAsnLeuValAlaGlyIyGlyIyThrIleAsnGlyAsnGlyGlnVal	180
Dd	504	GATAGGTCTCAAAATTTAGTTGTTGGAGAGAGAGAACTATCAAGGCAATGGAACAATA	56

QY	181	TTTTTTPROSERSerCysLys111eAsnLysSerLeuProCysArgaPaLaProtrala	200
Db	564	TTGATGCCAAGTCTTCTGCATAATATATATACATGCATGCAGGATGCACCAAGGCC	623
QY	201	LeuThrPheTpaEncySLysAsnLeuLysValAsnAsnLeuLysSer	216
Db	624	CTTAACTTCTGCATTCGCAAAATTTTGAAGTGATATATCTAAAGAGT	672
RESULT 15			
LOCUS	BM536089		
DEFINITION	BM536089	703 bp	MRNA linear EST 20-FEB-2002
ACCESSION	U85789111	tomato breaker fruit Lycopersicon esculentum cDNA clone	
VERSION	CLBG71124	5' end, mRNA sequence.	
KEYWORDS	BM536089.1	GI:18815229	
SOURCE	EST.		
ORGANISM	Lycopersicon esculentum (tomato)		
REFERENCE	Lycopersicon esculentum		
AUTHORS	Bukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon. 1 (bases 1 to 703)		
TITLE	Alcala,J., Vrebalov,J., White,R., Vision,T., Karamycheva,S.A., Ronning,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.		
JOURNAL	Generation of ESTs from tomato fruit tissue, breaker stage (2002)		
COMMENT	Unpublished (2002)		
CONTACT	Clemson University Genomics Institute		
ADDRESS	100 Jordan Hall, Clemson, SC 29634, USA		
EMAIL	Email: http://www.genome.clemson.edu/orders/index.html		
INSTITUTE	This clone is available through the Clemson University Genomics Institute		
Seq primer: T3.			
Location/Qualifiers			
1..703			
/organism="Lycopersicon esculentum"			
/mol_type="mRNA"			
/cultivar="TA496"			
/db_xref="taxon:4081"			
/clone="CLBG71124"			
/tissue_type="Pericarp"			
/dev_stage="breaker"			
/lab_host="SOLR"			
/clone_lib="tomato breaker fruit"			
/note="Vector: pBluescriptSKmCudaPc; Site 1: EcoRI; Site 2: XhoI; supplier: Boyce Thompson Institute; sequencing: The Institute for Genomic Research. Fruit were harvested at the breaker stage (first sign of lycopen accumulation on the blossom end of fruit). Fruit were cut in half and the seeds and locules were discarded prior to freezing the pericarp."			
Alignment Scores:			
Pred. No.:	3,25e-117	Length:	703
Score:	1110.00	Matches:	217
Percent Similarity:	99.09%	Conservative:	0
Best Local Similarity:	99.09%	Mismatches:	0
Query Match:	46.44%	Indels:	2
DB:	4	Gaps:	0
US-10-691-374-2 (1-457) x BM536089 (1-703)			
QY	1	MeVal11eglnhrgaSenSer11euleulellellellephealAserser11eser	20
Db	45	ATGTTATCCAAAGAGATAGTATTCCTTCATCATATATTTGGCTTCATCAATTTCA	104
QY	21	ThrCysArgSerAsnVal11leapapAsnLeuPheLysGlnVal1rAspAsn11leu	40
Db	105	ACTTGAGAGCAATGATTTATGATGACATTTATTCAAACAGTTATGATATATTTCTT	164

QY 121 GluGlnGluPheAlaHisAspPheGlnAlaTyrLeuSerTyrLeuSerLysAsnIleGlu 60

Db 165 GAAACAAGAAATTTGGCTCAAGATTTTCAAGCTTATCTTTCTATTATGACCAAAAATATTGAA 222

QY 61 SerAsnAsnAsnIleAspLysValAspLysAsnGlyIleLysValIleAsnValLeuSer 80

Db 225 AGCAACAATATATTATGACAAAGTTGATTAATAATGGATTTAANGATTAATGACTTAAGC 284

QY 81 PheGlyAlaLysGlyAspGlyLysThrTyrAspAsnIleAlaPheGlnAlaTyrAsn 100

Db 285 TTGGAGCTAAAGGCTGATGGAAACAACTATGATTAATTATGCAATTTGACCAAGCATGGAAT 344

QY 101 GluAlaCysSerSerArgThrProValGlnPheValProLysAsnLysAsnTyrLeu 120

Db 345 GAAACATGTTCACTCTGAAACACCTGTGAATTTGTGTTCTTAATAAACAAAGATTAATCTT 404

QY 121 LeuArgGlnIleThrPheSerGlyProCysArgSerSerIleSerValLysIlePheGly 140

Db 405 CTCAAGCAAAATCACTTTTCAGAGTCCAGACGAATCTCTATTCAGTAAAGATTTTGGGA 464

QY 141 SerLeuGlnAlaSerSerLysIleSerAspTyrLysAspArgIleLeuTyrIleAlaPhe 160

Db 465 TCCTTAGAAGCATCTAGTAAATTTCAACATTCAAAGATGAAGAGCTTTGGAAATTCGTT 524

QY 160 eAspSerValGlnAsnLeuValVal-GlyGlyGlyGlyThrIleAsnGlyAsnGlyGlnIle 180

Db 525 TGAATAGTTCCTCAAAATTTTGAATGTTGGAGAGAGAGGAACTATCAATGGCAATGGCAAG 584

QY 180 AlaTyrTrpProSerSerCysLysIleAsnLysSerLeuProCysArgAspAlaProThr 200

Db 585 TATGGTGGCCAAAGTTCTTGGCAAAATTAATAATCACTGCATCCAGAGATGACCAACGCG 644

QY 200 ILeuThrPheThrAsnCysLysAsnLeuLysValAsnAsnLeuLysSerLys 217

Db 645 CCTTAACCTTCCTGGATTCGAAAAATTTGAAAGGAATATATCTAAAGAGTTAA 697

```
Search completed: March 5, 2005, 01:15:42
Job time : 3535 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 5, 2005, 00:17:31 ; Search time 661 Seconds
(without alignments)

4102.523 Million cell updates/sec

Title: US-10-691-374-2

Perfect score: 2390

Sequence: 2 MYOINSLILIIIFASIS.....VTPHCTSLISDEALINNY 457

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3401638 seqs, 2966923429 residues

Total number of Hits satisfying chosen parameters: 10803276

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame_p2n.model -DEV=xlp
-O=/cgn2_1/USPTO/seqop1_p/US10691374/rnact 28022005 120708 21090/app_query.fasta.1.647
-DB=Published Applications NA -QMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=prc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10691374 @CGN 1.1 723 @rnact 28022005 120708 21090
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOPE=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
6: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09D_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
20: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
21: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
22: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the total being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1820	76.2	5822	10 US-09-924-197-1	Sequence 1, Appl1
2	1506	63.0	7456	18 US-10-691-374-1	Sequence 1, Appl1
3	1498	62.7	7456	18 US-10-691-374-47	Sequence 47, Appl1
4	1498	62.7	7456	18 US-10-691-374-49	Sequence 49, Appl1
5	1127	47.2	1621	18 US-10-739-930-530	Sequence 530, Appl1
6	1113	46.6	1631	13 US-10-151-668-1	Sequence 1, Appl1
7	1103	46.2	1657	18 US-10-787-958-31	Sequence 1548, Appl1
8	1103	46.2	1707	18 US-10-739-930-1548	Sequence 24607, A
9	1011	42.3	1187	18 US-10-437-963-24608	Sequence 24608, A
10	1010	42.3	1359	18 US-10-437-963-28561	Sequence 28561, A
11	988.5	41.4	1335	18 US-10-437-963-28561	Sequence 618, Appl1
12	971.5	40.6	1168	17 US-10-260-228-618	Sequence 33752, A
13	940	39.3	1372	18 US-10-425-115-33752	Sequence 30261, A
14	870	36.4	1308	17 US-10-425-114-30261	Sequence 52002, A
15	808	33.8	1573	18 US-10-437-963-32002	Sequence 31774, A
16	785	32.8	1866	18 US-10-437-963-31774	Sequence 82276, A
17	770	32.2	1623	18 US-10-425-115-82276	Sequence 70206, A
18	739	30.9	1723	18 US-10-437-963-70206	Sequence 2789, Appl1
19	734	30.7	1512	18 US-10-437-963-2789	Sequence 9445, Appl1
20	733.5	30.7	1822	18 US-10-739-930-2789	Sequence 92601, A
21	728	30.5	1631	17 US-10-425-114-9445	Sequence 318, Appl1
22	728	30.5	1984	17 US-10-425-559-92601	Sequence 141807, A
23	727	30.4	1673	17 US-10-362-091-3	Sequence 1144, Appl1
24	726.5	30.4	2007	18 US-10-425-115-3138	Sequence 2520, Appl1
25	723.5	30.3	1479	17 US-10-424-599-141807	Sequence 52557, A
26	709.5	29.7	1733	18 US-10-424-599-141807	Sequence 3023, Appl1
27	702.5	29.4	1185	9 US-09-938-842A-2520	Sequence 129068, A
28	702.5	29.4	1185	11 US-09-938-842A-2520	Sequence 26774, A
29	701.5	29.4	1203	18 US-10-437-963-52557	Sequence 2788, A
30	700	29.3	1884	17 US-10-425-114-3023	Sequence 56617, A
31	700	29.3	1919	18 US-10-425-115-129068	Sequence 59667, A
32	699	29.2	1833	17 US-10-425-114-26774	Sequence 43683, A
33	691.5	28.9	1644	18 US-10-739-930-2788	Sequence 13631, A
34	682	28.5	2219	18 US-10-437-963-56617	Sequence 33478, A
35	681	28.5	2174	18 US-10-437-963-59667	Sequence 10070, A
36	679.5	28.4	1374	18 US-10-437-963-33683	Sequence 4296, Appl1
37	666	27.9	1182	18 US-10-437-963-33683	Sequence 17989, A
38	660	27.6	1599	17 US-10-424-599-33478	Sequence 14146, A
39	656.5	27.5	1943	18 US-10-425-115-10070	Sequence 95670, A
40	655	27.4	1388	17 US-10-425-114-4296	Sequence 44782, A
41	655	27.4	1473	17 US-10-425-114-17989	Sequence 14060, A
42	654.5	27.4	1455	17 US-10-425-114-14146	
43	654.5	27.4	1472	17 US-10-424-599-35670	
44	654.5	27.4	1701	18 US-10-425-115-44782	
45	653.5	27.3	1312	17 US-10-425-114-14060	

ALIGNMENTS

RESULT 1
US-09-924-197-1
Sequence 1, Application US/09924197
GENERAL INFORMATION:
APPLICANT: Guttererson, Neal
TITLE OF INVENTION: Repeated Methods of Gene Silencing Using Inverted
FILE REFERENCE: 012176-01081005
CURRENT FILING DATE: 2001-08-07
PRIOR APPLICATION NUMBER: 60/225,508
NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 5822
TYPE: DNA
ORGANISM: Agrobacterium tumefaciens

US-09-924-197-1

Alignment Scores:

Pred. No.:	4,282-200	Length:	5822
Score:	1820.00	Matches:	345
Percent Similarity:	100.00%	Conservative:	1
Best local Similarity:	99.71%	Mismatches:	0
Query Match:	76.15%	Indels:	0
DB:	10	Gaps:	0

US-10-691-374-2 (1-457) X US-09-924-197-1 (1-5822)

QY	112	ValValProLysAsnLysAsnLysAsnLysLeuLeuLysGlnIleThrPheSerGlyProCysArg	131
Db	1387	ATGGTTCCTAAAAACAAGAAATTCCTCTCCAGCAAAATCACTTTTCAGGTCATGCAGA	1444
QY	132	SerSerIleSerValLysIlePheGlySerLeuGlnAlaSerSerLysIleSerAspTyr	151
Db	1447	TCTTCTTTTCAGTAAGAATTTTGGATTCCTTGAAGAGCATCAATAATTTTCAGACTAC	1506
QY	152	LysAspArgArgLeuTrpIleAlaPheAspSerValGlnAsnLeuValValGlyGlyGly	171
Db	1507	AAAGATAGAAGCTTGGATTCCTTTTGATAGCTTCAAAATTTAGTTCTTGGAGAGGA	1566
QY	172	GlyThrIleAsnGlyAsnGlyGlnValTrpTrpProSerSerCysLysIleAsnLysSer	191
Db	1567	GGAACTATCAATGGCAATGACACAGATAGTGGCCAAAGTTCTTGCAAAATTAATTAACA	1626
QY	192	LeuProCysArgAspAlaProThrAlaLeuThrPheTrpAsnCysLysAsnLeuLysVal	211
Db	1627	CTGCCATGACGGGATCACCAACGGCTTAACCTTCGAAATTCGAAAAATTTGAAAGTG	1686
QY	212	AsnAsnLeuLysSerLysAsnAlaGlnGlnIleHisIleLysPheGlySerCysThrAsn	231
Db	1687	AATTAATCTAAGAGTAAAAATGACACAAATTCATATCAATTTGAGTACGACTAAT	1746
QY	232	ValValAlaSerAsnLeuMetIleAsnAlaSerAlaLysSerProAsnThrAspGlyVal	251
Db	1747	GTTGTAGCTTCAAAATTTGATGTGATCATGCTTCACGAAAGAGCCCAATATCTGATGAGCTC	1806
QY	252	HisValSerAsnThrGlnTyrIleGlnIleSerAspTrpIleIleGlyThrGlyAspAsp	271
Db	1807	CATGATCAAAATCTCAATATATTCAAATATGATGATCTATTATTTGGAACAGGAGAT	1866
QY	272	CysIleSerIleValSerGlySerGlnAsnValGlnAlaThrAsnIleThrCysGlyPro	291
Db	1867	TGTATTTCAATGTTTCTCGAATCTCAAAATGTGCAGGCCACAAATATTACTTGTGGTCCA	1926
QY	292	GlyHisGlyIleSerIleGlySerLeuGlySerGlyAsnSerGlnAlaTyrValSerAsn	311
Db	1927	GGTCACTGGATTAAGATTAATGGAAGCTTGAAGATCTGGAAATTCAGAAAGCTTATGTCTAAT	1986
QY	312	ValThrValAsnGlnAlaLysIleIleIleGlyAlaGluAsnGlyValAlaGlyIleValThrTrp	331
Db	1987	GTTACTGTAATGAAGCCAAATATTAATCGGTGCCAAAATGGAGTTAGATCAACACTTGG	2046
QY	332	GlnGlyGlySerGlyGlnAlaSerAsnIleLysPheLeuAsnValGluMetGlnAspVal	351
Db	2047	CAGGAGGATCTGGACAAAGCTAGCAACATCAAAATTTCTGAATGTGGAAATGCAAGAGCTT	2106
QY	352	LysTyrTrpIleIleIleAspGlnAsnTyrCysAspArgValGluProCysIleGlnGln	371
Db	2107	AAGATATCCCAATTAATTAAGCAAAACTATTTGATGAGTGAACATGATATACACAG	2166
QY	372	PheSerAlaValGlnValLysAsnValValTyrGluAsnIleLysGlyThrSerAlaThr	391
Db	2167	TTTTCAAGCATTTCAAGTGAATAATGTGTGTATGAGATATCAAGGGCCACAAAGTGCACA	2226
QY	392	LysValAlaIleLysPheAspCysSerThrAsnPheProCysGlnGlyIleIleMetGln	411
Db	2227	AAAGTGGCCATATAATTTGATTTGACAGACCAAACTTTTCATGTGAAGAAATTAATAAGAG	2286
QY	412	AsnIleAsnLeuValGlyGluSerGlyLysProSerGlnAlaThrCysLysAsnValHis	431

D _b	2287	AATTTAAATTAGTAGGGGAAAAGTGAAAAACATCAAGGCTCAGTCCAAAATGTGCAT	2346
Q _y	432	PheaaanaaaiaagiurhisvaltlnrProHieCyseTherleuguiileserGiuaapgiu	451
D _b	2347	TTTTACAATCTGAACACTGTTAACCCCACTGCACITTCATAGAATAATTTCAGAGATGAA	2406
Q _y	452	AlaleuleurfyrantYr	457
D _b	2407	GCTCTTTTGATATAATTAT	2424

RESULT 2

```

US-10-691-374-1
/ Sequence 1, Application US/10691374
/ Publication No. US20040230322A1
/ GENERAL INFORMATION:
/ APPLICANT: McCallum, Claire
/ APPLICANT: Slade, Ann J.
/ APPLICANT: Colbert, Trent
/ APPLICANT: Koanf, Vic
/ APPLICANT: Anawah Inc.
/ TITLE OF INVENTION: Tomatoes Having Reduced Polygalacturonase Activity Caused by Non-
/ TITLE OF INVENTION: Transgenic Mutations in the Polygalacturonase Gene
/ FILE REFERENCE: MHB 02-276
/ CURRENT APPLICATION NUMBER: US/10/691,374
/ NUMBER OF SEQ ID NOS: 50
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 1
/ LENGTH: 7456
/ TYPE: DNA
/ ORGANISM: Lycopersicon esculentum
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1479)..(1757)
/ OTHER INFORMATION:
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (2416)..(2547)
/ OTHER INFORMATION:
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (3327)..(3491)
/ OTHER INFORMATION:
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (3696)..(3716)
/ OTHER INFORMATION:
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (4260)..(4467)
/ OTHER INFORMATION:
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (4567)..(4648)
/ OTHER INFORMATION:
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (5602)..(5710)
/ OTHER INFORMATION:
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (6139)..(6255)
/ OTHER INFORMATION:
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (6788)..(7045)
/ OTHER INFORMATION:
US-10-691-374-1

Alignment Scores:
Pred. No.: 2.48e-163 Length: 7456
Score: 1506.00 Matches: 454

```

Percent Similarity: 24.43% Conservative: 0
Best Local Similarity: 24.43% Mismatches: 3
Query Match: 63.01% Indels: 1404
DB: 18 Gaps: 8
US-10-691-374-2 (1-457) x US-10-691-374-1 (1-7456)

QY 1 MetValIleGlnArgAsnSerIleLeuLeuIleIleIlePheAlaSerSerIleSer 20
DB 1479 ATGGTATTCMAAGAGATAGTATTCCTTCATCATATTAATTTTCCTCATCAATTTCA 1538
QY 21 ThrCysArgSerAsnValIleAspAspAsnLeuPheValIleValIleAsnIleLeu 40
DB 1538 ACTTGAGACGCAATGTATATGATGCAATTAATTAACAAGTTATGATATATTTCTT 1598
QY 41 GluGlnGluPheAlaHisAspPheGlnAlaTyrLeuSerTyrLeuSerTyrLeuSer 60
DB 1599 GAACAAGATTTGCTCATATTTTCAAGCTTATCTTTCTTATTTGACAAAATATTTGA 1658
QY 61 SerAsnAsnAsnIleAspIleValIleAspIleValIleValIleValIleAsnValIle 80
DB 1659 AGCAACAAATAAATGACAAAGTTGATTAATAAATGGGATTAAGTATGATGATCTTAC 1718
QY 81 PheGlyAlaIleValGlyAspGlyLeuThrTyrAspAsnIle----- 93
DB 1719 TTGGAGCTAAGGGTATGAGAAACATATGATATATTTGATATTTAAATTTGGAA 1778
QY 93 ----- 93
DB 1779 TATATTTGCGGGATGAAAATGATAGAGATATTAAGATTAATTTGGAGATGAAAAGTT 1838
QY 93 ----- 93
DB 1839 ATATTTTATAAAGTAAATAATATTTCTCGTTTTAGTATTAAGGTGAAAATGAGTTTC 1898
QY 93 ----- 93
DB 1899 TCGTTAAGCGAGAAAAGCTATTTTCCATGCTAATCTTATTTTCTTACTTTAATAA 1958
QY 93 ----- 93
DB 1959 CGTCATAGATTTTGTCTACTACAGAAATAGACACTTATTTATGATGATTTAGTGTGCA 2018
QY 93 ----- 93
DB 2019 AAAGAAATGATAGTAATTTTGTCTAATATTAATCAATTTCTATATGATATATTTTC 2078
QY 93 ----- 93
DB 2079 AACCAAAATACAAAGCTAATCCATTAAGTGGGCTTAGAAATAAAGATAGTTCTAT 2138
QY 93 ----- 93
DB 2139 TCAATTTCTAACCTTATTTAATTTTAGTGAACCTCGACAAAACGAACAACGTATTC 2198
QY 93 ----- 93
DB 2199 AAACCTTTATATTCGGAATTCGAGACCAACATATGACCAACCTCACATGATATAGT 2258
QY 93 ----- 93
DB 2259 CCTAATATATTAATTTTCTAATAAAATATCTTCAATCTACATATGCAAAATTTGAAA 2318
QY 93 ----- 93
DB 2319 ATGACTTTATCTCTATCGAACACATATACAGAGTTCTTTTAAGAATTTACACACTACAT 2378
QY 94 -----AlaPheGlnGlnAlaTyrPheGln 101
DB 2379 TTGGATATGTTTCTTATCGGTAAATTAATCTTTCAAGCAATTTGACACAGCATGATATA 2438
QY 101 uAlaCysSerSerArgThrProValGlnPheValValProValAsnIleValTyrLeuLeu 121

DB 2439 AGCATGTTCACTAGAACACCTGTTCAATTTTGTTGTTCTCTAATAAACAAAGATTAATCTCT 2498
QY 121 uLysGlnIleThrPheSerGlyProCysArgSerSerIleSerValIle----- 137
DB 2499 CAAGCAATACCTTTTCAGTGCATGACATCTTCTATTTCCAGT-AAAGTTAGCATAT 2557
QY 137 ----- 137
DB 2558 TGATATTTATATTCCTCTTTGTTAGCAATATATATCTGTTTANGACAAAATTTAAGAA 2617
QY 137 ----- 137
DB 2618 AGTAAATCAAGATAGATAAACAATGAATTTTGTCACTAATTTAGCGGATTTAGTGAGAA 2677
QY 137 ----- 137
DB 2678 TTATCAAAATGTATGTAGTATGACCAACTTAGCTATGAATTAGCTAGTGAAGAATT 2737
QY 137 ----- 137
DB 2738 TGATGTAATCTATTTTCTTATATATGATGATAGTTCAAACTAAAGACTGCAAAAGAA 2857
QY 137 ----- 137
DB 2798 TTAATATGCTTAATTAATATATGATGATAGTTCAAACTAAAGACTGCAAAAGAA 2857
QY 137 ----- 137
DB 2858 AATAAGAAAGAAATATTTATTTTAAATAATTAATAAAGAAAATATGAGAAATTAATTT 2917
QY 137 ----- 137
DB 2918 CAAAGCGAAGATATTACATATCTATGAGGATTAAGATTAATATATATGTAAGAAA 2977
QY 137 ----- 137
DB 2978 CAGCACTACATATCTAATAAAGTCTCAATAATGATATTAATAAATAGTGTATAGCAA 3037
QY 137 ----- 137
DB 3038 CAGTTATCCCTACAAAACCTTTTGTGGGTAGATGATGATCAGAGGTGTTCCAGACTCT 3097
QY 137 ----- 137
DB 3098 TCGTTAAAAAATGTTTCTTAATAAAGTTTGAAGAATGTTATATGATGAATAATAT 3157
QY 137 ----- 137
DB 3158 GAAAGAAAACATATCAATTTAATAAATTAATAAGTATCAAAAGTAAACGAATTAACATA 3217
QY 137 ----- 137
DB 3218 GGAATATATCTATTAATGAATAATTTAGTGCTTTGTTAACAATATCTAGTTATATTC 3277
QY 138 -----IlePheGlyIle 141
DB 3278 ATGTGTTCTTAATTTCCCTCTTATTTTGTGAATTAATCAATGACAGATTTTGGATTC 3337
QY 141 rLeuGlnAlaSerSerIleSerAspTyrIleAspArgIleTyrIleAlaPheAs 161
DB 3338 CTTAGAAGCATCTAGTAATAATTTTCAGACTACAAAGATAGAGGCTTTGGATTCCTTTGA 3397
QY 161 pSerValGlnAsnLeuValIleGlyIleGlyIleThrIleAsnGlyAsnGlyGlnValIle 181
DB 3398 TAGTGTTCAAAATTTAGTTGTGGAGAGAGAACTATCAATGCAATGACAAAGTATAG 3457
QY 181 pThrProSerSerCysIleValIleAsnIleSerLeu----- 192
DB 3458 GTGGCCAAAGTCTTGCAAAATAAATTAATCACTGTAATTTTAATACTTGTCTATATAGT 3517
QY 192 ----- 192
DB 3518 TTACGCTATGTTGCTCGAATTTCTTAACCTGTTCTAAGATATTAATATTTGAAGAA 3577

Alignment Scores:

Pred. No.:	2,12e-162	Length:	7456
Score:	1498.00	Matches:	453
Percent Similarity:	24.38%	Conservative:	0
Beac Local Similarity:	24.38%	Mismatches:	4
Query Match:	62.68%	Indels:	1404
DB:	18	Gaps:	8

US-10-691-374-2 (1-457) x US-10-691-374-47 (1-7456)

```

QY      1 MetValIleGlnArgAsnSerIleLeuLeuIleIlePheAlaSerSerIleSer 20
Db      1479 ATGGTATCCAAAGGAATGATTCCTCTTCATTAATTTTTCCTTCATTTCA 1538
QY      21 ThrCysArgSerAspValIleAspAspLeuPheIleGlnValTyrAspAsnIleLeu 40
Db      1539 ACTTGTAGAGCAAGCAAGTTATGATGACAAATTTATCAACAAGTTATGATTAATTTCT 1598
QY      41 GluGlnIlePheAlaHisAspPheGlnAlaTyrLeuSerTyrLeuSerTyrAsnIleGlu 60
Db      1599 GAACAGAGATTTGCTCATGATTTTCAGCTTATCTTTCTTTTGAAGCAAAATATTGAA 1658
QY      61 SerAsnAsnAsnIleAspIleValAspIleValIleValIleValIleValIleVal 80
Db      1659 AGCAACATTAATATGACAGGTTGATTAATAAATGGATTAAGGATTAATGACTTAC 1718
QY      81 PheGlyAlaIleValAspGlyIleValTyrAspAsnIle----- 93
Db      1719 TTTGGAGCTAAGGGGATGATGAAACATATGATTAATTTGTAAGTATTAAATATTGAA 1778
QY      93 ----- 93
Db      1779 TATATTTGGGGATGAAATGATAGAAATATAGAAATTTTGGAAAGATGAAAGTT 1838
QY      93 ----- 93
Db      1839 ATATTTTATAAGTGAATAATTAATTTTCCTGTTTAAATTAAGTGAATAAGTTTC 1898
QY      93 ----- 93
Db      1899 TCGTTAAGCAGGAAAGACTATTTCCATGTTAAGTATTTTCTTACTTTAATAA 1958
QY      93 ----- 93
Db      1959 CGTATAGTATTTGCTATCTACTCAAGATAAGACACTTATTAATGATTAAGTGTCTGA 2018
QY      93 ----- 93
Db      2019 AAGAATATGATGATTAATTTGCTTATATTAATCTATCAATTTCTTAATATTAATTTTC 2078
QY      93 ----- 93
Db      2079 AACCAAAATTAACAAGCGTATTCATTAAGTGGGCTTGAATAAGGTAAGTTCTAT 2138
QY      93 ----- 93
Db      2139 TCAATTTCTAACCTTATTTAATTTTAAAGTGAACCTCGACAAAACGAACAAAGTATTC 2198
QY      93 ----- 93
Db      2199 AAACCTTTATATTCGGAATTCGAGCAACATATGAACAACCTCACACATGATATAGT 2258
QY      93 ----- 93
Db      2259 CCTAATATATTAATTTTCTTAATAAATATCTTCAATCTACCATTAATGAATATTGAAA 2318
QY      93 ----- 93
Db      2319 ATGACTTTATTCATGACACATTAATCAAGAGTTCTTTTAAGATTTTACCATCAT 2378
QY      94 ----- AlaPheGlnIleAlaTyrPheGlnIle 101
Db      2379 TTGGATGTTCTTATGCTGTTAAATTAATCTTTGAGGCAATTTGAGCAACATGGAATGA 2438

```

```

QY      101 uAlaCysSerSerArgThrProValGlnPheValIleProIleAsnIleAsnTyrLeuLe 121
Db      2439 AGCATGTTCACTTAGAACAACCTGTTCAATTTTGTTCTTAATAAACAAATTAATCTTCT 2498
QY      121 uLysGlnIleThrPheSerGlyProCysArgSerSerIleSerValys----- 137
Db      2499 CAAGCAATACACCTTTTCAGGTCATGCAAGATCTTCTATTTCAAGT-AAAGTTAGCAATAT 2557
QY      137 ----- 137
Db      2558 TGAATTTATATTCCTTTGTTAGCAATAATATATCTGTTTATAGCAAAATTAAGAA 2617
QY      137 ----- 137
Db      2618 AGTAATCAAGAATAGATAAACAATGAATTTTCGTCATAATTTAGCGGATTAAGAGAA 2677
QY      137 ----- 137
Db      2678 TATCAAAATGTTATGTTAGTATGACAACTTAGATGAATTAAGTGAAGAAATT 2737
QY      137 ----- 137
Db      2738 TGATGCTAATTCATTTTCTTTTGTAGATGAATAATTTGAAACACATGATTAATTA 2797
QY      137 ----- 137
Db      2798 TTAATTAATGCTTAATTAATATGTCATGCAATGACTTCAACTAAGAACTGTCAAAAGAA 2857
QY      137 ----- 137
Db      2858 AATAAGAAAGAAATTAATTTTAAATAATTAATAAGAAAGAAATATGAGAAATAATT 2917
QY      137 ----- 137
Db      2918 CAAGCGAAGAGATTAACATTAATCTATGGGATTAAGATTAATTAATTAATTAAGAAA 2977
QY      137 ----- 137
Db      2978 CAGACTACACATCTTAATTAAGTCTCATTAATGATATTAATAAATATGTTAAGCAA 3037
QY      137 ----- 137
Db      3038 CAGTTATCCCTAAGAAAACCTTTTGGGGATGATGATCATCAGAGTTGTTCCAGACTCT 3097
QY      137 ----- 137
Db      3098 TGCTTAATAAATATGTTTCTTAATTAAGTTTGAAGAATGTTATATGATGAATAATAT 3157
QY      137 ----- 137
Db      3158 GAAGAAAACATATCAATTTTAATAATAATAATAAGTAATCAAGTAACGAATAAATAA 3217
QY      137 ----- 137
Db      3218 GGAATATACATATAATGAATAATTAAGTGGCTTTTCGTTAATCAATATCTAGTTATTC 3277
QY      138 ----- 137
Db      3278 ATTTGTTCTTAATTTCCCTCTTAATTTTGTGAATAATTAATGACAGATTTTGGATC 3337
QY      141 rLeuGlnIleAspSerIleSerAspTyrIleAspArgIleLeuPheAlaPheAs 161
Db      3338 CTTAGAGCATCTAATAAATTTTCAGATCAAAAGATGAAGGCTTTGGAATGCTTTTGA 3397
QY      161 pSerValGlnAsnLeuValIleGlyGlyIleThrIleAsnGlyAsnGlyGlnValTr 181
Db      3398 TAGTGTTCAAAATTAATTTGTTGGGAGGAGAACTATCAATGCGCATATAGACAAGTATG 3457
QY      181 pTyrProSerSerCysIleAsnIleAsnIleSerLeu----- 192
Db      3458 GTGGCAAGTTCTTGCAAAATTAATAATCACTGTAATTAATTAACCTTGCTTAATAGT 3517

```

192 ----- 192
Db 3518 TTTACGCTATGTTGCTCGAATCTTTAACTGTTCTAAGATATATATATTTGAAGGA 3577
Qy 192 ----- 192
Db 3578 GGTGCACAAATGCATCATTTTAAAGATTCCGACCAATATAGTTTATGTAATCTA 3637
Qy 193 ----- 193
Db 3638 ATTTTCAGAGCATCTTGCCCTGTACTGATCATTTGTAACCTTTTTTCTTCATGCAGCC 3697
Qy 193 OcyarigaspA1Prothf----- 199
Db 3698 ATGCAGGAGTGCAACGCGTACGTTAATGTCATTGATTGATTAATAAAAAAGCCCTA 3757
Qy 199 ----- 199
Db 3758 AATATATTTGAATTTTAAATGAAAGTTATATAATCTTAACCTTGGCAGACCTAT 3817
Qy 199 ----- 199
Db 3818 TAACCCCTGCATATTAATAGTATTTTAAAGATATAAAGTGTGTTGAACAACA 3877
Qy 199 ----- 199
Db 3878 AATTTAGATATTCAAAACTATTGAAAAATTACTATAAATGCAATTTTTCATATCA 3937
Qy 199 ----- 199
Db 3938 ATATGATTTAAAAATATTAAGTTCTTATGATTGATTCATAAATAAATAATCATG 3997
Qy 199 ----- 199
Db 3998 ACAACAATAGTAGACGAGAAAGTATATAACAATACCTCTCAAGTAGAATCGATTGT 4057
Qy 199 ----- 199
Db 4058 ACACACTCAAAACCTACGTTTCTTCGATTTATATTCCTATTTCTTTAATAGTAA 4117
Qy 199 ----- 199
Db 4118 TCMAAGGCTATAGTTCTGTCAAAATCTATACATTGGAACCTATCTTTGACGCTGT 4177
Qy 199 ----- 199
Db 4178 ACATTGAGATGTTGAACAATGAAATGATTAATTAATTTGATTTAATATTAATA 4237
Qy 200 ----- 212
Db 4238 ACTAATATGTTTAAATTTTCAGGCTTAACCTTCGGAATTCGAAAAATTTGAAGTGAA 4297
Qy 212 nasenleuYseRlyAsnA1aglInI1ehIseRlyPhaGluSeRlyThrAsnVa 232
Db 4298 TATATTAAGAGTAAAAATGCACACAAATTCATATCAATTTGAGTCAATGACATAT 4357
Qy 232 lVal1aseRasleuMetI1eAsnA1aseR1alySeRProAsnThrAspGlyValIhI 252
Db 4358 TGTACTCTCAAAATTTGATGATCAATGCTTCAGCAAAAGCCCAAAATACTGATGAGTCCA 4417
Qy 252 sValseRasnThrgInI1eRlyseRphrI1eI1eGlyThr----- 268
Db 4418 TGTATCAATACTCAATATATTCAAATATCTGATATTAATTTGAAACGGTTATTTAT 4477
Qy 268 ----- 268
Db 4478 TTAATTTTATTTATCCAAATTTAATAGAAAAAAAAGAGATTTTATTTGATACATA 4537
Qy 269 ----- 279
Db 4538 AATTAATTAATTTTATTTTATTTTATTAAGTATGATTTGATTTCAATTTGTTCTGAGTC 4597
Qy 279 rclnMenValGlnA1eThrAsnI1eThrCyseGlyProGlyhI1eGlyI1eSeR----- 296

Db 4598 TCAAAATGACGGCCACAAATATTTACTTGTGTCCAGCTCATGTATAG-GTACTCTA 4656
Qy 296 ----- 296
Db 4657 TTTTACAATATACTGTTTCCATTTTCTATTTCAATAAAGTAGATATATATTA 4716
Qy 296 ----- 296
Db 4717 TTACTTAATTCCTTAAATTAATTTATTTGGCAAAATTTTCTCTGTCTTAAT 4776
Qy 296 ----- 296
Db 4777 GACTTAGCAATATATTAGGCCGTTTGATGGCGGAATAAAGCAGCTTTAAAAAGTA 4836
Qy 296 ----- 296
Db 4837 CTTTAAAAAGTGTGAACCTTATTTTAAATAAGCAGTTATCGGTTGGATPAAAAAGTGC 4896
Qy 296 ----- 296
Db 4897 TGAAGTTTATGTCAACGTCGAAAAAGGAAAAATGAAAGAAATGTTAGGTTATA 4956
Qy 296 ----- 296
Db 4957 TGCGTATTTGTATAAAAATATTAAGCACAAAAAGTAAATGTGTCACTTAAACA 5016
Qy 296 ----- 296
Db 5017 ACTTAAGCTACCTACCCCTACCCAGCTTTAATCTTTGCTTAAATAAGTTTTT 5076
Qy 296 ----- 296
Db 5077 TTTTAAAACTTAAATAAGTTGTTTGTAGTATTTGCCAAAGACTTAATATGCAAAACC 5136
Qy 296 ----- 296
Db 5137 AGCTTTAAGTCAGTTGACGACGCTTTTAACTGAGCCAAACAGGCTTTAAATGTCTG 5196
Qy 296 ----- 296
Db 5197 CTTAGATGTCTATATATATTTAGCTTTTGTGAAGTAGATATATATCTTAAGTTCAA 5256
Qy 296 ----- 296
Db 5257 CATAAAAATACATGCTTTAACAATAGCAATATAGTTAATCAAAAAAGCAAAATGATGA 5316
Qy 296 ----- 296
Db 5317 TTTTGCGAATTTGATTTATTCACAGAAAAAGGATAGTCAAGGTATTCATGAAT 5376
Qy 296 ----- 296
Db 5377 TGAAGATATCATTAAGACTTAATAATTAAGAATCAATATTAAGGAGTCAAAAAATGTAT 5436
Qy 296 ----- 296
Db 5437 TACCTTAATAAATCTATTCATTTGATTAATAATTAATACTAATTAAGAGTGTATA 5496
Qy 296 ----- 296
Db 5497 ATCTAATAAACATGCATTAATTTGACGAATGTGTTTGTGACTATATCTTTCT 5556
Qy 297 ----- 301
Db 5557 GAAATATGCTATATTTTCTCTTTTATTTTTCATGAGATTAATGGAAGCTTAG 5616
Qy 301 ySeRlyAsnSerGlyA1eYrValseRasnVal1eThrVal1aenGlyA1eVal1eI1eG1 321
Db 5617 ATCTGAAATTCAGAGCTTAATGTCTTAATGTACTGTAAATGAAGCCAAATATTAATCG 5676
Qy 321 yA1aGluenGlyVal1eRgI1eYrThrTrpGln----- 332

Db 5677 TGCCGAAATGAGTTAGATCAAGCTTGGA-GGTACCTCCCCCCCCCCCCCCCC 5735
QY 332 ----- 332
Db 5736 CAGAGCCCATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 5795
QY 332 ----- 332
Db 5796 AATTGATTCATGTTGAAATTTAATTTGATTAAGATTAAGTATTTAATTTAATTTA 5855
QY 332 ----- 332
Db 5856 GTTATATAGAAAAAAATTTTCAAGACTTCAGATTATTTGATCTAGTAAAT 5915
QY 332 ----- 332
Db 5916 GTTGTCTTTGTTGAAAGTTTGTTATTCAGTTTGGGTCAATGATTAAACCAACTTA 5975
QY 332 ----- 332
Db 5976 TAATGAAAAAGGGGTGCAACGGCGCCCACTAGTCTAGTATCAATAGAAATCTCAC 6035
QY 332 ----- 332
Db 6036 GTCTGTTTATTCAGATGAGCGTCTGTGTAATGTTAAATTAATTAATTAACA 6095
QY 333 ----- 338
Db 6096 TGTATTAAGCATTATATAATTAATGTTATTAATTAATTAATTAATTAATTAATTA 6155
QY 338 ----- 358
Db 6156 TAGGACATCAAAATTTCTGATGTGAAATGCAAGAGTAAATATCCCATATTTAAGA 6215
QY 358 ----- 370
Db 6216 CCAAACTATGTGATGAGTTGACATGATTAACAAGTAATTTTAAATTAACAAG 6275
QY 370 ----- 370
Db 6276 AATTATTAATTTTAACTCTTAATCACTTACATTAATAAATTGAGATCTT 6335
QY 370 ----- 370
Db 6336 TTCACTAGTAGTAACCTTTTGAATGATTTTAAATGATTAATTAATTTCTTTTA 6395
QY 370 ----- 370
Db 6396 TTTTCTTAAATTAATGATCTTTTGACTATGCTTAATAATCTTTTAAGTAAC 6455
QY 370 ----- 370
Db 6456 GAATATCATAGAAAAAATGTAATGATTAATTAATTTTAAATTAATTAAGAGTT 6515
QY 370 ----- 370
Db 6516 ATCTATATTTTGTATGTAAGTGAAGAAATATACATAAATTAAGATTTCAAAATTTA 6575
QY 370 ----- 370
Db 6576 AATAATCTAAATAATTTGATTAATCAATCATGACTACAGAAAGAAATACGAGATT 6635
QY 370 ----- 370
Db 6636 CTCAAAATATAGATGAGTCTTTCAAAACGAATCAAAATTAATGATTAATCTTTAA 6695
QY 370 ----- 370
Db 6696 AAGAGACATTTGGTAATAATGTAATAATCATTAAGTTAATAAATAATTTCAACTGA 6755
QY 371 ----- 381
Db 6756 TCTCTCAGATTAATTAATCACTTTGTTCCAGTTTCAAGATTCAGTAAATAATGTGG 6815

QY 381 aTyrGluAsnIleValThrSerAlaThrIleValAlaIleValPheAspCysSer 401
Db 6816 TGTATGAAATATCAAGGCGACAAAGTGCATTAATAATTTGATTCAGACA 6875
QY 401 hrAsnProCysGluGlyIleIleMetGluAsnIleAsnIleValGlyIleSerGly 421
Db 6876 CAACCTTCCATGTGAGGAATTTATATGAGAAATTAATTAATTTAGTACGAGAAAGTGA 6935
QY 421 yProSerGluAlaThrCysIleValIleIlePheAsnAlaGluIleValThrPro 441
Db 6936 AACCATCAAGGCTACGTGCAAAATGTCATTTAACAATGCTGAACATGTTACACCAC 6995
QY 441 iCySerThrLeuGluIleSerGluAspGluAlaLeuLeuTyrAsnTyr 457
Db 6996 ACTGCATCTCACTAAGAAATTTTCAGAGATGAAGCTTTTGTATTAATTAAT 7045

RESULT 4
US-10-691-374-49
Sequence 49, Application US/10691374
Publication No. US20040250322A1
GENERAL INFORMATION:
APPLICANT: McCallum, Claire
APPLICANT: Slade, Ann J.
APPLICANT: Colbert, Trent
APPLICANT: Knauft, Vic
APPLICANT: Anawah Inc.
TITLE OF INVENTION: Tomatoes Having Reduced Polygalacturonase Activity Caused by Non-
FILE REFERENCE: MBH 02-276
CURRENT FILING DATE: 2003-10-22
NUMBER OF SEQ ID NOS: 50
SOFTWARE: PatentIn version 3.1
SEQ ID NO 49
LENGTH: 7456
TYPE: DNA
ORGANISM: Lycopersicon esculentum
FEATURE:
NAME/KEY: CDS
LOCATION: (1479)..(1757)
OTHER INFORMATION:
FEATURE:
NAME/KEY: CDS
LOCATION: (2416)..(2547)
OTHER INFORMATION:
FEATURE:
NAME/KEY: CDS
LOCATION: (3327)..(3491)
OTHER INFORMATION:
FEATURE:
NAME/KEY: CDS
LOCATION: (3696)..(3716)
OTHER INFORMATION:
FEATURE:
NAME/KEY: CDS
LOCATION: (4260)..(4467)
OTHER INFORMATION:
FEATURE:
NAME/KEY: CDS
LOCATION: (4567)..(4648)
OTHER INFORMATION:
FEATURE:
NAME/KEY: CDS
LOCATION: (5602)..(5710)
OTHER INFORMATION:
FEATURE:
NAME/KEY: CDS
LOCATION: (6139)..(6255)
OTHER INFORMATION:
FEATURE:
NAME/KEY: CDS
LOCATION: (6788)..(7045)

OTHER INFORMATION:

US-10-691-374-49

Alignment Scores:

Pred. No.:	2,12e-162	Length:	7456
Score:	1498.00	Matches:	453
Percent Similarity:	24.38%	Conservative:	0
Best Local Similarity:	24.38%	Mismatches:	4
Query Match:	62.68%	Indels:	1404
DB:	18	Gaps:	8

US-10-691-374-2 (1-457) x US-10-691-374-49 (1-7456)

```
QY      1 MetValIIIGlnArgAsnSerIIleuLeuLeuIIleIIlePheAlaSerSerIIleSer 20
Db      1479 ATGGTTATCCAAAGAAATAGATATTCCTTCCTCATATATATTTGGCTTCATCAATTTCA 1538

QY      21 ThrCysArgSerAsnValIIleAspAspAsnLeuPheValIleProIysAsnIleLeu 40
Db      1539 ACTTGTAGAGCAAGCTTATTTGATGACATTTATTCAAACAAAGTTATGATATATTTCTT 1598

QY      41 GlnGlnIlePheAlaIleAspPheGlnAlaIleLeuSerIleLeuSerIleGln 60
Db      1599 GAACAAGATTTGCTCATGATTTTCAAGCTTATCTTTCTTATTTGACAAATAATTTGA 1658

QY      61 SerAsnAsnAsnIIleAspIleValIleAspIleAsnGlyIleValIleAsnValIleuSer 80
Db      1659 AGCAACAATATATTTGACAAAGGTTGATTAATAAATGGATTAAGATATATGACTTAC 1718

QY      81 PheGlyAlaIleValIleAspGlyIleAspGlyIleAspPheAsnIle----- 93
Db      1719 TTTGGAGCTAAGGGTGTGTGAGAAACATATGATATATTTGTAAATTTAAATTTGAA 1778

QY      93 ----- 93
Db      1779 TATATTTGGGGATGAAATATGATAGATATTAAGAAATTTTGGAAAGATGAAAGTT 1838

QY      93 ----- 93
Db      1839 ATATTTTAAAGTAAATAATATTTTCTCGTTTTAGTATTAAGGTGAAATGAGTTTC 1898

QY      93 ----- 93
Db      1899 TCGTTAAGCGAAGAAAGCTATTTCCATGCTACTGATTTTTTTTTTACTTTTAATA 1958

QY      93 ----- 93
Db      1959 CGTCATAGTATTTGCTATCTCAAGAAATTAAGACTTATTTATTTGATTTAGTCTCGA 2018

QY      93 ----- 93
Db      2019 AAAAAGATTTGATGATTAATTTTGGCTTAATAATACTATTCATATTTCTATATGATTTTTC 2078

QY      93 ----- 93
Db      2079 AATCAAAATTAACAAGCGTATTCATTAAGTGGGCTTGAATAAAGTAGTTCTAT 2138

QY      93 ----- 93
Db      2139 TCAATTTCTTAACCTTATTTATTTAGTGAACCTCGAATAAAGAAACAAACGATATTC 2198

QY      93 ----- 93
Db      2199 AAACCTTTATATTTGGAATTCGAGACCAACATATGAACAACCTCAACATCATATAGT 2258

QY      93 ----- 93
Db      2259 CCTATATATATATTTTCTAATAAATAATCTTCAATCTACATATTAATAATTTGAAA 2318

QY      93 ----- 93
Db      2319 ATGACTTTATCTATCGAACACATATCAAGATTTCTTTAAGATTTTACCACTACAT 2378
```

```
QY      94 -----AlaPheGlnGlnAlaIleTPAsnGly 101
Db      2379 TTGGTATGTTCTTATCGGTAAATAATATCTTTCAGGCACTTTGAGCAAGCTGCAATTA 2438

QY      101 uAlaCysSerSerArgThrProValGlnPheValIleProIysAsnIleValIleuLeu 121
Db      2439 AGCATGTTCACTTACAAACCGTTCATATTTGTGTTCTTAATAAACAAGAAATTAATCTTCT 2498

QY      121 uIleGlnIleThrPheSerGlyProCysArgSerSerIleSerValIle----- 137
Db      2499 CAAGCAATTCACCTTTTCAGGTCATGACAGATCTTCTATTTCAGT-AAAGTTAGCATAT 2557

QY      137 ----- 137
Db      2558 TGATTTATTTATCTCTTTGTTAGCAATATATTTCTGGTTATGACAAATTTAAGA 2617

QY      137 ----- 137
Db      2618 AGTATCAAGATAGATAAACAATGAATTTGTCACATAATTTAGCGATAGTGAGAA 2677

QY      137 ----- 137
Db      2678 TTATCAAAATGTTATGTTAGCTATGAGCAACTTACCTATGAAATTAAGTGAAGAGTT 2737

QY      137 ----- 137
Db      2738 TGATCTAATTTCTATTTTTTTTTTTGTAGAGTAAAGATATTTGAAACACATGATTAATTA 2797

QY      137 ----- 137
Db      2798 TTAATTTATGTTAATTAATATGATGATGATGTTCAACTAAAGAACTGCAAAAGAA 2857

QY      137 ----- 137
Db      2858 AATAAGAAAGAAATTTATTTTAAATAATTAATAAAGAAATATGAGAAATTAAT 2917

QY      137 ----- 137
Db      2918 CAAGCGAAGAGTTTACATATCTATGAGGATTAAGAGATTTATATGTAAGAAA 2977

QY      137 ----- 137
Db      2978 CAGCACTACATATCTAATAAAGTCTAATAAGATATTAATAAATAAGTGTAAAGAA 3037

QY      137 ----- 137
Db      3038 CAGTTATCCCTACAAAACTTTTGGGGTAGATCATCCAGAGGTGTTTCCAGACTCT 3097

QY      137 ----- 137
Db      3098 TGCCTTAATAAATGTTTTTTCTAATAATAAGTTGAAGAAATGTTATATGATGAATAAT 3157

QY      137 ----- 137
Db      3158 GAAGAAAAACATATCATATTTAATAAATAAAGTATCAAAAGTAAAGAAATTAACAATA 3217

QY      137 ----- 137
Db      3218 GGAATATATCTATTAATGAATAATTTAGTGCTTTGCTTAACATATCTTAGTTATTC 3277

QY      138 -----1lePheGlyIle 141
Db      3278 ATTGTTCTTTAATTTCCCTTCTTATTTTGTGAATTTACTTAATGACAGATTTTGGATC 3337

QY      141 IleuGlnIleSerSerIleSerAspIleValIleAspArgArgLeuTPIIleAlaPheAsn 161
Db      3338 CTTAGAACATCTAGTAATAATTTGACACTACAAAGATGAAGGCTTTGGATTTGCTTTGA 3397

QY      161 pSerValGlnAsnLeuValIleGlyGlyGlyThrIleAsnGlyAsnGlyValIle 181
Db      3398 TAGTGTTCAAATTTAGTGTGGAGAGAGAACTATCATATGCAATGACCAAGTATG 3457

QY      181 pThrProSerSerCysIleAsnIleSerLeu----- 192
```

Db 3458 GTGGCCAGTCTTGGAAAATAAATCACTGTAATTTTATACTTGGCTTAAAGT 3517
QY 192 -----
Db 3518 TTTACGCTATGTTGTCGAATCTTTAACTTGTCTAAAGATATTATATTTGAAGA 3577
QY 192 -----
Db 3578 GGATGCAAAATGATCATATTTTAGAGATTCGACCAATATTAGTTTATGTAATCTA 3637
QY 193 -----Pr 193
Db 3638 AATTTCAGAGCATCTTGGCTTGTACTGATCATTTGTACCTTTTCTTCATGACAGCC 3697
QY 193 OCSATGAPALAPROTH-----
Db 3698 ATGCGAGGATGCAACGATCGTTAAATTCATTTGATTGATTAATAAAAAAAGCCTA 3757
QY 199 -----
Db 3758 AAATATATTTGAATTTTAATTGAAGTTATATATATTTCTTAACCTTGGCAGAGCCTAT 3817
QY 199 -----
Db 3818 TACCCCTTGCATTTAATATGATGATTTTAAAGATATAAAGTTTGTAGTGAACAA 3877
QY 199 -----
Db 3878 AATTAGATATTCAAAAAATTTGAAATTTACTATTAATTCGAATTTTTCGATATCA 3937
QY 199 -----
Db 3938 ATATGATTAAAAATATTAGTTAAAGTTCTTAGATTGATCTGAAAAATTAATCATG 3997
QY 199 -----
Db 3998 ACAACAATAGTAGACGGAAGATATATACATACCTTTCAGTAGATCGATTGT 4057
QY 199 -----
Db 4058 ACAACACCTCAAAACCTACGTTTCTTGATTTATATTTCCATTTCTTTAATAGTA 4117
QY 199 -----
Db 4118 TCAAAGCTATTAGTTCTGCAAAATCTATACATTGAAACCTATCTTGAAGCTGCT 4177
QY 199 -----
Db 4178 ACATTCGAGATCGTGAACATGATGATATTATTAACCTTGTATTTAAATTTAA 4237
QY 200 -----AlaLeuThrPheTyrPancCysIysAsnLeuIysValAs 212
Db 4238 ACTAATATTGTTAATTTTCAGGCTTAACTTCGGAATTCAGAAAAATTGAAGTGA 4297
QY 212 nAsnLeuIysSerIysAsnAlaGlnGlnIleHisIleIysPheGlySerCysThrAsnVa 232
Db 4298 TATATCAAGAGTAAAAATGACACAAATTCATATCAATTTGATCATCACTAAAGT 4357
QY 232 IValIAserAsnLeuMetIleAsnAlaSerAlaIysSerProAsnThrAspGlyValHi 252
Db 4358 TGTACCTCAAAATTTGATGATCAATGCTTCAGCAAAAGGCCAAATACTGATGAGTCCA 4417
QY 252 sValSerAsnThrGlnIyrIleGlnIleSerAspThrIleIleGlyThr----- 268
Db 4418 AGTACCAAAATCTCAATATATTCAAAATCTGATATCTATTTATGAAACAGGTTATTTAT 4477
QY 268 -----
Db 4478 TTAATTTTATTTATCAATTTAATTGAAAAAAGAGTATTTTATTTGATTAACTA 4537
QY 269 -----GlyAspAspCysIleSerIleValSerGlyse 279

Db 4538 AATTATTAATTTTAAATTTTATTTTATAGTAGATGATTTATTTCAATTTGCTTGATTC 4597
QY 279 rGlnAsnValGlnAlaThrAsnIleThrCysGlyProGlyHisGlyIleSer----- 296
Db 4598 TCAAAATGTCAGGCGCCAAATATTTACTTGTGTCAGGTCAATGATATAG- GTA CTCTA 4656
QY 296 -----
Db 4657 TTTTACAAATATATCTTGTTCATTTTCTCTATTTCAATTAAGTAGATATATATTA 4716
QY 296 -----
Db 4717 TTAATTAAATCTTTAAATTAATTTATGCAAAATTTTCTCTGTCTTAATGTTAAT 4776
QY 296 -----
Db 4777 GACTTAGCACAATATTAGGCGCGTTGATGGCGAATAAAGCAGCTTTAAAAAGTA 4836
QY 296 -----
Db 4837 CTTTAAAAAGTGTGAACCTTATTTTAAATTAAGCATTTATCGGTTTGATTAAGTGC 4896
QY 296 -----
Db 4897 TGAAGTTGTATGTCAAACGTGAAAAAGGAAAAATGAGAGAAAGAAATGTAGGTTATA 4956
QY 296 -----
Db 4957 TGGGTTATTGTATAAAAATATTATAGCACAAGAAATAAAAAGTGCTCACTTAAAAACA 5016
QY 296 -----
Db 5017 ACTATTAAGTACCCTAACCCACCTTTTAACCTTTTGCTTAAATTAAGTTTCTT 5076
QY 296 -----
Db 5077 TTTTAAAACTTAAATAAGTTGTTTGTAGTATTTGCCAAAGAGCTAATATATGCAAAAACC 5136
QY 296 -----
Db 5137 AGCTTTAAGTCAGTTTGACACAGCTTTTAAGTCGAGCAACAGAGCTTTAAATGCTG 5196
QY 296 -----
Db 5197 CTTAGATGTCATATATATTTAGAGCTTTTGTGAAGTATATATATCTTAAGTTCA 5256
QY 296 -----
Db 5257 CATTAATAACATGCTTTAACAATAGCACATATAGTTATCAAAAAGCAAAATGATGATTA 5316
QY 296 -----
Db 5317 TTTTGGAAATTTGATTATTCACAGAAAAAGGATAGTTCAAAGTATCAATTCATGAT 5376
QY 296 -----
Db 5377 TGAAGATATCATTAAGACTAAATTTAGAAGATCAATAATTGAGGATCAAAAAATGTTAT 5436
QY 296 -----
Db 5437 TACCTTATTAATAATCTATTCATTTTCATATTAATTAACATAATTAAGAGTGTATA 5496
QY 296 -----
Db 5497 ATCTAATAAACAATGATATATATGACGAATGTGTTTGGTACCTTAATCTTTCT 5556
QY 297 -----IleGlySerLeuG 301
Db 5557 GAATATTTGCTATTTTCTCTTTTATTTTTCATGATTAATCTTAATGGAACCTTAGG 5616
QY 301 ySerGlyAsnSerGlyAlaTyrValIserAsnValThrValAsnGlnAlaIysIleIleG 321
Db 5617 ATCTGAAATTCAGAGCTTATGATGCTTAATGTTACTGTAAATGAAGCCAAATTTATCGG 5676

```

QY 321 YAAAGLUBANGIYVALARGILEYTHRTPLN----- 332
Db 5677 TCCCCAAAATGAGTAGATCAAGACTTGCA-GGTACCTCCCCCCCCCCCCCCCC 5735
QY 332 ----- 332
Db 5736 CAGAGCCCATTTTTTAATTTTTTTAAATTTTATTCGAATATCAATATTAAGATTAA 5795
QY 332 ----- 332
Db 5796 ATTGATTTTCATGTTGAAATTTATATTTGATTAAGATATTTTACTAGCTTCTAT 5855
QY 332 ----- 332
Db 5856 GTTATATGAAAAAAAATGTTCCAGAACCTTCAATATTATCTGCTACTAAGTAAAT 5915
QY 332 ----- 332
Db 5916 GTGTTCCTTTGTTTGAAGTTGGTTTATCCAGTTTGGGTCATGATTAAACCAACTTA 5975
QY 332 ----- 332
Db 5976 TAAATGAAAAAGGGCTGCACAGCGCCGACCTAGTCTAGTATCATATGAAAGATCTAC 6035
QY 332 ----- 332
Db 6036 GTCTGTTTATTCAGATGAGCGTTCTTGTTGATGTTAATTAATTAATTAATTAACA 6095
QY 333 ----- 338
Db 6096 TGTAAATTAAGCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6155
QY 338 ASERANILEYSPHELEUASENVALGLUWETGLNAPVALYSTRPROILEIIELEAS 358
Db 6156 TAGCAACATCAAAATTTCTGAATGTGGAATGCAAGCGTTAAGTATCCCATATATATA 6215
QY 358 PGJNABENTYCYBAPBAPGVALGLUPROCYSLIEGLN----- 370
Db 6216 CCAAAAATAATGTGATCGAGTTGAACCATGTATACAAAGGTAAATTTTATTAAAGAAC 6275
QY 370 ----- 370
Db 6276 AATTTATTAATTTTATTAATCTTCTTAATACCTTACATCAATTAATAAATTGAGATTCT 6335
QY 370 ----- 370
Db 6336 TTCACCTAGTAGTAACCTTTTGAATAGATTTTGAATAGATTTTGAATAGATTTCTTTTA 6395
QY 370 ----- 370
Db 6396 TTTTCTCTTAATTTATGATCTTTTGAATAGTCTTAATAATCTTGTTAAAGTAACT 6455
QY 370 ----- 370
Db 6456 GAATATCATAGAAAAAATGTTAGATTATTAATCTAAATTTTATTAAATTTATTAGACGTT 6515
QY 370 ----- 370
Db 6516 ATCTAATATTTTGTATGTAAAGTTGAAACATATATCAATTAACATTTAGATTCAATTTTA 6575
QY 370 ----- 370
Db 6576 ATTAATATCTAAATATTTGATGCAATCAATCATGACTACAAACGAATACATGACAGATT 6635
QY 370 ----- 370
Db 6636 CTCAAACATATATGATGAGACTCATTTCAAAACGAATCAAAATATATGTAAGATATCTTAA 6695
QY 370 ----- 370
Db 6696 AAGAGCATTTGGTAAATTAAGTAAATAATCATTTAAGTTATAAAAAAATTTCTAACTGA 6755

```

```

QY 371 -----GlnPheSerAlaValGlnValLeuAsnVal 381
Db 6756 TCTCAGCATTATTTAATCACTTGTTCACATTTTCACATTTCAAGGAAATATGCG 6815
QY 381 ALTYRGLUASNILEYSGLYTHRSEALATHLYVALALALEYSPHEAPCYSERT 401
Db 6816 TGTATGAGAAATATCAAGGCCACAGTGCACAAAGTGCCCAATTAATTTGATTGACGA 6875
QY 401 hAsnPheProCysGluGlyIleIleMetGluAsnIleAsnLeuValGlyGluSerGly 421
Db 6876 CAACCTTTCATGATAGGAATTAATATGAGAAATTAATTAATTTAGTACGGGAAAGTGA 6935
QY 421 YAPROSEGLUATHRCELYSASENVALHISPHASASNNLAGLUNISVALTHPRDH 441
Db 6936 AACCATCAGAGGCTACGGCAAAAATGTCATTTTAAACATGCTGACATGTTACACAC 6995
QY 441 IECYETHRSELEUGLUILESEGLUAPSGLUALEUENTYRANTRY 457
Db 6996 ACTGCACCTTCATAGAAATTTTCAGAGGATGAAGGCTCTTTGTRATATATAT 7045

RESULT 5
US-10-739-930-530
; Sequence 530, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; FILE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53377) B
; CURRENT APPLICATION NUMBER: US/10/739,930
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 530
; LENGTH: 1621
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: Clone ID: ARATH-23APR03-CLUSTER21779_1
US-10-739-930-530

Alignment Scores:
Pred. No.: 2,866-120 Length: 1621
Score: 1127.00 Matches: 223
Percent Similarity: 65.33% Conservative: 71
Best Local Similarity: 49.56% Mismatches: 116
Query Match: 47.15% Indels: 40
DB: 18 Gaps: 6

US-10-691-374-2 (1-457) x US-10-739-930-530 (1-1621)
QY 7 SerIleLeuLeuLeuIleIleIlePheAlaSerSerIleSerThrCysArgSerAsnVal 26
Db 124 GCTGTTTCTTAATGCGTCTTTTGAATGCTCGTGTGTGTAAGCTTTGAGTACAGCAAGTT 183
QY 27 IleAsp-----AspAsnLeuPheLeuGlnVal 35
Db 184 GATGATGATATGTCATGATGAGATGAGATTTGAATCCGATGACTTGTCTCAAGCTCAAC 243
QY 36 TyrAspAsnIleLeuGluGlnGluIlePheAlaHisAspPheGlnAlaTyrLeuSerTyrLeu 55
Db 244 AACGATGATGTTCTT-----AGCTTGATTAAGCTCAGAC 276
QY 56 SerIleAsnIleGluSerAsnAsnAsnIleAspIleValIleAspIleAsnGlyIleLeuVal 75
Db 277 GAACCACTTGTGAAGCATCA-----ACC 300
QY 76 IleAsnValIleuSerPheGlyValIleYsgIYAspGlyLeuThrTyrAspAsnIleAlaPhe 95
Db 301 GTTAGTGTTTCAAACTTCGAGCCCAAGAGATGAGAAATGATGATATCTCAGGCTTC 360
QY 96 GluGlnAlaTyrAsnGluIleCysSerSerAspGlyThrProValGlnPheValIleProLeu 115
Db 361 AAGAAAGCATGGAAGAAAGCAATGTCACAAATGAGATTACTTCTTGGTCTCTAAA 420

```

```

QY 116 AsnLysAsnTyrLeuLeuLysGlnIleThrPheSerGlyProCysArgSerIleSer 135
Db 421 GGAAGAGCTATCTCTTAAAGTCTACTCGATTAGAGCCCAAGCAATCTTACGTAAAC 480
QY 136 ValLysIlePheGlySerLeuGlnAlaSerSerLysIleSerAspTyrLysAspArg 155
Db 481 TTTCAGATCCTTAGGCTTATTCAGATCTACGAAACGTTCCAGATTTCAGAAACAAAAC 540
QY 156 LeuTyrIleAlaPheAspSerValGlnAsnLeuValGlyGly-----GlyThr 173
Db 541 CATTGGCTTATCTTAAAGAGCGTTAACATATCATCATGACGAGGCTCGACGGGAATT 600
QY 174 IleAsnGlyAsnGlyGlnValTyrTyrProSerSerCysLysIleAsnLysSerLeuPro 193
Db 601 ATTAAATGGCAACGGAAGAAACCTGGTGGCAAGCTCATGCAAAATGCAAAATCTAAGCA 660
QY 194 CysArgAspAlaProThrAlaLeuThrPheTyrAsnCysLysAsnLysValAsnAsn 213
Db 661 TGCAAAAGCTCCAAAGCGCTCTTACTTATCAATTAAAGATTGAATGTGAAGAT 720
QY 214 LeuLysSerLysAsnAlaGlnGlnIleHisIleLysPheGlySerCysThrAsnVal 233
Db 721 CTGAGGGTGAAAAAGCCGAGAGATTCAAGTTTCATTTGAGAAATGCACAAAGTTGAA 780
QY 234 AlaSerAsnLeuMetIleAsnAlaSerAlaLysSerProAsnThrAspGlyValHisVal 253
Db 781 GTTACTTAATGTTGAGATCATCTCGGCGATAGTCCCAACAGATGATCATCATTC 840
QY 254 SerAsnThrGlnTyrIleGlnIleSerAspThrIleLeuGlyThrGlyAspAspCysIle 273
Db 841 ACTAATACTCAAAACATTCCAGCTCCAACTCAAGATTCGGAACAGGTGATGATTGATA 900
QY 274 SerIleValSerGlySerGlnAsnValGlnAlaThrAsnIleThrCysGlyProGlyHis 293
Db 901 TCCTATGAGATGAGCAAGCAAAATCTTCAATCTTGAATTAAGTTCGCGCCCGCTCAC 960
QY 294 GlyIleSerIleGlySerLeuGlySerGlyAsnSerGlnAlaTyrValSerAsnValThr 313
Db 961 GGATTCAGCATTTGGGAGCTTGGGAGAGACACAAATTCAAACCTTATGTCGGAATTAAT 1020
QY 314 ValAsnGlnAlaLysIleIleGlyAlaGluAsnGlyValArgIleLysThrTrpGlnGly 333
Db 1021 GTGATGAGTGTCTAAGTCTCTGAGAGTGCACATGAGATTAGATTAAAGCTTATCGGGA 1080
QY 334 GlySerGlyGlnAlaSerAsnIleLysPheLeuAsnValGlnMetGlnAspValLysTyr 353
Db 1081 GGAATCAGGAAGCTCCAGAAACATTTAAATTTCAAAATATTCGATGGAATAACCTCAAGAT 1140
QY 354 ProIleIleIleAspGlnAsnTyrCysAspArgValGluProCysIleGlnGlnPheSer 373
Db 1141 CCGATCATATATCCAGACAGACTACTGCGACAG-----GACAAATGCGAAAGACCAAGATCG 1197
QY 374 AlaValGlnValLysAsnValValTyrGluAsnIleLysGlyThrSerAlaThrLysVal 393
Db 1198 GCGATGCAAGTGAAGAAACCTTGTGTACAGAAACATATCTGGTACAGACGCTACCGAGTGTG 1257
QY 394 AlaIleLysPheAspCysSerThrAsnPheProCysGluGlyIleIleMetGlnAsnIle 413
Db 1258 GCGATTAAGCTGAATTCGAGCGAGAAAGTATCATATCAGGATTTGCTTGAAGAAAGCTG 1317
QY 414 AsnLeuValGlyGlySerGlyLysPheSerGlnAlaThrCysLysAsnValHisPheAsn 433
Db 1318 AAAATTAAGGA-----GGAACAGCTTCTGCAAAATGCGCAATGTTAAA 1362
QY 434 AsnAlaGluHisValThrProHisCysThr 443
Db 1363 AATCAAGGACCGCTTCTCTAAATGCTCT 1392

```

RESULT 6
 US-10-151-668-1
 : Sequence 1, Application US/10151668
 : Publication No. US20020184660A1

```

: GENERAL INFORMATION:
: APPLICANT: UNUSKOV, Peter
: APPLICANT: CHILD, Robin
: APPLICANT: VAN ONCKELIN, Henri
: APPLICANT: PRINSEN, Els
: APPLICANT: BORKHARDT, Bernard
: APPLICANT: SANDER, Lijli
: APPLICANT: PETERSEN, Morten
: APPLICANT: BUNDGARD, POUlsen, Gert
: APPLICANT: BOTTERMAN, Johan
: TITLE OR INVENTION: Seed Shattering
: FILE REFERENCE: 2121-0138P
: CURRENT APPLICATION NUMBER: US/10/151,668
: PRIOR FILING DATE: 2002-05-21
: PRIOR APPLICATION NUMBER: US/09/051,239
: PRIOR FILING DATE: 1998-09-28
: PRIOR APPLICATION NUMBER: PCT/EP96/04313
: PRIOR FILING DATE: 1996-10-04
: PRIOR APPLICATION NUMBER: EP 95 402241.4
: PRIOR FILING DATE: 1995-10-06
: PRIOR APPLICATION NUMBER: EP 95 203328.0
: NUMBER OF SEQ ID NOS: 14
: SOFTWARE: Patencin Ver. 2.1
: SEQ ID NO 1
: LENGTH: 1631
: TYPE: DNA
: ORGANISM: Brassica napus
: FEATURE:
: OTHER INFORMATION: Location 95-163 = region encoding the presumed
: OTHER INFORMATION: endo-PG signal peptide.
: OTHER INFORMATION: Location 884-900 = region of the endo-PG cDNA
: OTHER INFORMATION: corresponding to oligonucleotide PG3
: OTHER INFORMATION: Location 1059-1073 = region of the endo-PG cDNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (95)-(1393)
: OTHER INFORMATION: Location 1229-1245 = region of the endo-PG cDNA
: OTHER INFORMATION: complementary to oligonucleotide PG5
: OTHER INFORMATION: Location 821-837 = region of endo-PG cDNA
: OTHER INFORMATION: corresponding to oligonucleotide PGI.
: NAME/KEY: unsure
: LOCATION: (1439)
: OTHER INFORMATION:
: OTHER INFORMATION: Strain cv. Topaz.
: US-10-151-668-1
:
: Alignment Scores:
: Pred. No.: 1,23e-118 Length: 1631
: Score: 1113.00 Matches: 223
: Percent Similarity: 65.32% Conservative: 69
: Best Local Similarity: 49.89% Mismatches: 121
: Query Match: 46.57% Indels: 34
: DB: 13 Gaps: 8
:
: US-10-691-374-2 (1-457) x US-10-151-668-1 (1-1631)
:
: QY 7 SerIleLeuLeuLeuIleIleIlePheAlaSerSerIleSerThrCysArgSerAsnVal 26
: Db 119 GCTGTTTCTTATGCGCTTTTGTATGCTCGCTGTCGCAAGCTTTGAGTACGAACGTA 178
:
: QY 27 IleAspAsnLeuLeuPheLysGlnValTyrAspAsnIleLeuGluGlnGluPheAlaHis 46
: Db 179 ---GATGATGGA-----TATGTCAT 196
:
: QY 47 AspPheGlnAlaTyrLeuSerTyrLeuSerLysAsnIleGluSerAsnAsnAsnIleAsp 66
: Db 197 GAAGATGGA-----AGCTTCGAATCCGATAGTTATCAAGCTCAACGACGACGAC 247
:
: QY 67 LysVal-----AspLysAsnGlyIleLysValIleAsnVal 78

```



```
Db 248 GACGTTCTTACCTTAAAAAGCTCTGATAGACCCACTACCGAATCATCTGTAAGT 307
Qy 79 LeuSerPheGlyAlaIleGlyAspGlyThrTrpAspAsnIleAlaPheGluGlnAla 98
Db 308 TCGAATCTTCGAGCCAAAGAGATCGAAGAAAACCGATGATATCTCAGGCTTCAAGAAAGCA 367
Qy 99 TrpAsnGluAlaCysSerSerArgThrProValGlnPheValValProIleAsnIleAsn 118
Db 368 TGGAGAAAGGCAATGTTCAACAAATGAGAGTACTACTTCTTAATTCCTTAAGAAAGAGCT 427
Qy 119 TyrIleuLeuGlyGlnIleThrPheSerGlyProCysArgSerSerIleSerValIle 138
Db 428 TATCTCTTAAGTCAATTAAGATTCAGAGCCCATGCAATCTTACGTAAGCTTCCAGATC 487
Qy 139 PheGlySerLeuGluAlaSerSerIleAspIleSerAspIleSerValIleSerValIle 157
Db 488 CTAGGCACTTTATTCAGCTTCTTCAAAACGATCGAATTACAGTAAGACAGAAACCACTGG 547
Qy 158 IleAlaPheAspSerValGlnAsnLeuValValGlyGlyGly-----GlyThrIleAsn 175
Db 548 CTATTTTGGAGAGCGTTAATTAATCTATCATCATGATGGCGGCTCGCGGAGATTGTGAT 607
Qy 176 GlyAsnGlyGlnValIleTrpTrpProSerSerCysIleAsnIleAsnIleSerIleProCysArg 195
Db 608 GCGAAGCGAAATATCTGGTGGCAAACTCATGCAAAATGACAAATCTTAAGCCATGACACA 667
Qy 196 AspAlaProThrAlaLeuThrPheTrpAsnCysIleAsnIleValIleAsnIleAsn 215
Db 668 AAGGCGCCAGCGCTTACTCTTACCAACCTAAAGAAATTTGAATGTAAGATCTGAGAA 727
Qy 216 SerIleAsnAlaGlnGlnIleIleIleIlePheGlySerCysThrAsnValValAlaSer 235
Db 728 GTGAGAAATGCAACACAGATTCAGATTGCAATTCGAAACAAATTTGGCGGTTAG 787
Qy 226 AsnIleuMetIleAsnAlaSerAlaIleIleIlePheAsnIleAspGlyValIleValSerAsn 255
Db 788 AATGTTAAGATCACTGCTCTCGGATGATCCCAACGAGATGATATCTGTTGCT 847
Qy 256 ThrGlnIleIleGlnIleSerAspThrIleIleGlyThrGlyAspAspCysIleSerIle 275
Db 848 ACTAAACCATTCGATCTCCAAATTCAGACATTTGGGACGATGATGTAATTCATT 907
Qy 276 ValSerGlySerGlnAsnValGlnAlaIleAsnIleThrCysGlyProGlyIleGlyIle 295
Db 908 GAGGATGGATCCCAAAATGTTCAATTCATTAATCTTGGCGCCCGCTGATGGAGAC 967
Qy 296 SerIleGlySerLeuGlySerGlyAsnSerGluAlaIleValSerAsnValThrValAsn 315
Db 968 AGCATTTGAGAGCTTGGGAGATGACATTCCAAGCTTATGTAATCGGAAATGATGAT 1027
Qy 316 GlnAlaIleIleIleGlyAlaGlnIleAsnGlyValArgIleIleThrTrpGlnIleGlySer 335
Db 1028 GGTGCTAGCGCTCTCGAGACGTGACATGAGATGAGATCAAGACTTACCAAGGAGGATCA 1087
Qy 336 GlyIleGlnAlaSerAsnIleIlePheLeuAsnValGlnIleAspValIleSerIleProIle 355
Db 1088 GGAATCTGCTTAAGAAATTAATTCCAAAACATTCGATGATGATATGTCAAAGATCGGAC 1147
Qy 356 IleIleAspGlnAsnIleCysAspArgValGlnIleProCysIleGlnIlePheSerAlaVal 375
Db 1148 ATTAATCGACAACTACTGCGACAAAG--GACAAATGCGAACACAGCAATCTGCGGTT 1204
Qy 376 GlnValIleAsnValValIleGlnIleAsnIleIleGlyThrSerAlaThrIleValAlaIle 395
Db 1205 CAAGTGAACAAATGCTGATATCAAGAAATTAAGGATGACAGCGCAACAGATGCGGATTA 1264
Qy 396 LysPheAspCysSerThrAsnPheProCysGlnIleIleIleMetGlnIleAsnIleAsn 415
Db 1265 ATGTTAATTCAGAGTGAATATCATCTCAAGGATTTGCTGATGATGATGAAACATC 1324
Qy 416 ValGlyGlySerGlyLysProSerGlnAlaIleThrCysIleAsnValIleIlePheAsnAla 435
Db 1325 AAGGAA-----GGAAGAGCTTCTGCGAAATGTCATGTTAAGATTA 1369
```

```
Qy 436 GlnIleValThrProIleCys 442
Db 1370 GGCACCTGTTCTCTTAATGC,1390

RESULT 7
US-10-787-958-31
; Sequence 31, Application US/10787958
; Publication No. US20040154053A1
; GENERAL INFORMATION:
; APPLICANT: Wyatt, Paul
; APPLICANT: Roberts, Jeremy A.
; APPLICANT: Whitelaw, Catherine
; TITLE OF INVENTION: Signal Transduction Protein Involved in Plant Defence
; FILE REFERENCE: 0623,0890000
; CURRENT APPLICATION NUMBER: US/10/787,958
; CURRENT FILING DATE: 2004-02-27
; PRIOR APPLICATION NUMBER: GB9806113.8
; PRIOR FILING DATE: 1998-03-20
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 31
; LENGTH: 1657
; TYPE: DNA
; ORGANISM: Brassica napus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (145) .. (1443)
US-10-787-958-31

Alignment Scores:
Pred. No.: 1,856-117 Length: 1657
Score: 1103.00 Matches: 221
Percent Similarity: 65.54% Conservative: 70
Best Local Similarity: 49.77% Mismatches: 125
Query Match: 46.15% Indels: 28
DB: Gaps: 7

US-10-691-374-2 (1-457) x US-10-787-958-31 (1-1657)
Qy 7 SerIleLeuLeuIleIleIlePheAlaSerSerIleSerThrCysArgSerAsnVal 26
Db 169 GGTATTTTCTTATGCGTTCTTTGATGCGTCTGCGCAAGCTTGAAGTGAACGTA 228
Qy 27 IleAsp-----AspAsnIlePheIleGlnValIleTrpAsnIleLeuGln 41
Db 229 GATGATGATATGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 282
Qy 42 GlnGlnPheAlaIleAspPheGlnAlaIleIleIleIleIleIleIleIleIleIleIle 61
Db 283 CTCAACAAAGCAAGCAAGCTTCTTACCTTGAAGCTCGATGACCCACTACCGAATCA 342
Qy 62 AsnAsnAsnIleAspIleValIleAspIleValIleValIleValIleValIleValIle 81
Db 343 TCA-----ACTGTTAGTCTTTCGAATCTC 366
Qy 82 GlyAlaIleGlyAspGlyLysThrTrpAspAsnIleAlaPheGlnIleAlaIleTrpAsnGln 101
Db 367 GAGGCAAAAGGATGAGAAACCGATGATATCTCAGGCTTTCAGAAACATGCAAGAAAG 426
Qy 102 AlaCysSerSerArgThrProValGlnPheValValProIleAsnIleAsnIleIleIle 121
Db 427 GATGTTCAACAAATGAGTGACTACTTCTTGAATTCCTTAAGGAGAAAGCTTATCTCTT 486
Qy 122 LysGlnIleIleThrPheSerGlyProCysArgSerSerIleSerValIlePheGlySer 141
Db 487 AAGTCTATTAGATTCAGAGGCCCATGCAAAATCATTAACGATTCAGATTCAGAGCACT 546
Qy 142 LeuGlnAlaSerSerIleIleSerAspIleIleIleIleIleIleIleIleIleIleIle 160
Db 547 TTATAGCTTTCACAAACGATCGGATTAACGTAATGACAAAGCACTGCTATTTTGG 606
Qy 161 AspSerValGlnAsnLeuValValGlyGlyGly-----GlyThrIleAsnGlyAsnGly 178
```

```

Db      607 GAGGACCTTAATATCTATCATCATGATGCGGCTCGCGGCTTGTATGACCAACGGA
Qy      179 GlnValTrpTrpProSerSerCysLysIleAsnLysSerLeuProCysArgAspIlePro
Db      667 AAAATCGGTGGCAAAATCATCATCAAAATCGAATATCAAGCATGCAAAAGCGCCA
Qy      199 ThrAlaLeuThrPheTrpPheCysLysAsnLeuLysValAsnAsnLeuLysSerLysAsn
Db      727 ACGGCTTACTCTCTACCAACTTAACAAATTTGAATGACAAATCTGAGAGTGAAGAAAT
Qy      219 AlaGlnGlnIleHisIleLysPheGluSerCysThrAsnValAlaLysSerLeuMet
Db      787 GCACACGACAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG
Qy      239 IleAsnAlaSerAlaLysSerProAsnThrAspGlyValHisValSerAsnThrGlnTyr
Db      847 ATCACTGCTCTCGGCATATGCTCCCAACAGATGATTCATCATCTGCTACTATAAAAC
Qy      259 IleGlnIleSerAspThrIleIleGlyThrGlyAspAspCysIleSerIleValSerGly
Db      907 ATTCGAATCTCAATTCAGACATTCGGACAGCTGATTCATTCATTCATTCATTCATTCATTC
Qy      279 SerGlnAsnValGlnAlaThrAsnIleThrCysGlyProGlyHisGlyIleSerIleGly
Db      967 TCGCAAAATGTTCAATCAATGATTTAACTTGGCGCCCGCTCATGCGATCAGCATTCGA
Qy      299 SerLeuGlySerGlyAsnSerGlyAlaLysValSerAsnValThrValAsnGlnAlaLys
Db      1027 AGCTTGGGGGATGCAATTCCAAACTTATATCGGAATTAATGATGATGATGATGATGATG
Qy      319 IleIleGlyAlaGlnAsnGlyValArgIleLysThrTrpGlnGlySerGlyGlnAla
Db      1087 CTCCTGAGACTGACAAATGAGTAAGATCAAGACTTACCAAGGAGGCTCAGAACTCCT
Qy      339 SerAsnIleLysPheLeuAsnValGlnMetGlnAspValLysTrpProIleIleLeuAsp
Db      1147 AAGAACATTAATTCCAAAACATTCGATGATTAATCAAGATTCGATCAATTAATCGAC
Qy      359 GlnAsnTrpCysAspArgValGluProCysIleGlnGlnPheSerAlaValGlnValLys
Db      1207 CAGAACCTACTGCGCAAG---GACAAATGCGAACCAAGAACTCGGTTCAAGTGAAC
Qy      379 AsnValValTrpGlnAsnIleLysGlyThrSerAlaThrLysValAlaIleLysPheAsp
Db      1264 AAGTGTGTATCGGAACATCAAGGTCAGCGCAAGCATGTCGATTAATGTTTAAAT
Qy      399 CysSerThrAsnPheProCysGlnGlyIleIleMetGlnAsnIleAsnLeuValGlyGlu
Db      1324 TGCAGGTGTGAATATCATGCGCAAGGATTTGCTTGAGATGATGATCACTAAAGA---
Qy      419 SerGlyLysPheSerGlnAlaThrCysLysAsnValHisPheAsnAsnAlaGlnHisVal
Db      1381 -----GGAAGACTTCTTGCAGAAATATGCAATGTAAGATTAAGACGACCGTT
Qy      439 ThrProHisCys 442
Db      1429 TCTCTTAATGC 1440

```

```

RESULT 8
US-10-739-930-1548
; Sequence 1548, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; FILE REFERENCE: 38-21(53377)B
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 1548
; LENGTH: 1707

```

```

; TYPE: DNA
; ORGANISM: Brassica napus
; FEATURE:
; OTHER INFORMATION: Clone ID: BRAVA-23APR03-CLUSTER5_1
US-10-739-930-1548

Alignment Scores:
Pred. No.: 1,95e-117 Length: 1707
Score: 1103.00 Matches: 221
Percent Similarity: 65.54% Conservative: 70
Best Local Similarity: 49.77% Mismatches: 125
Query Match: 46.15% Indels: 28
DB: 18 Gaps: 7

US-10-691-374-2 (1-457) x US-10-739-930-1548 (1-1707)
Qy      7 SerIleLeuLeuLeuIleIleIlePheAlaSerSerIleSerThrCysArgSerAsnVal
Db      169 GCATATTTCTTATWGCGTTCTTTTGAATGCTCGCTGCGCAAGCTTGAAGTGAAGCACTTA
Qy      27 IleAsp-----AspAsnLeuPheLysGlnValTyrAspAsnIleLeuGlu
Db      229 GATGATGATATGTCATGAAGATGAAGACTTCGAAAC-----GATAGTTTATATCAAG
Qy      42 GlnGluPheAlaHisAspPheGlnAlaTyrLeuSerTyrLeuSerLysAsnIleGluSer
Db      283 CTCACACACGACGACGACGCTTACCTTGAAAAGCTCCGATAGACCCACCTACCGAATCA
Qy      62 AsnAsnAsnIleAspLysValAspLysAsnGlyIleLysValIleAsnValLeuSerPhe
Db      343 TCA-----ACTGTATGTTTCGAACCTTC
Qy      82 GlnAlaLysGlyAspGlyLysThrTyrAspAsnIleAlaPheGlnGlnAlaTrpAsnGlu
Db      367 GGAACAAAAGGTGATGAAAAACCGATATCTACAGCTTTCAGAAAGCATGGAAGAAG
Qy      102 AlaCysSerSerArgThrProValGlnPheValAlaProLysAsnLysAsnTyrLeuLeu
Db      427 GCATGTTCAACAAATGAGAGTACTACTTCTTATCTTAAGGAAAGCACTTATCTCCTT
Qy      122 LysGlnIleThrPheSerGlyProCysArgSerSerIleSerValLysIlePheGlySer
Db      487 AAGCTATTAATGATTCAGAGCCCATCAATCTTAATCGTACGTTCCAGATTCCTAGGCATC
Qy      142 LeuGlnAlaSerSerLysIleSerAspTyr---LysAspArgArgLeuTrpIleAlaPhe
Db      547 TTATCAGCTTCAACAAAGATCGATTCAGTATTAAGACAAACCACTGCTTATTTTG
Qy      161 AspSerValGlnAsnLeuValAlaGlyGlyGly-----GlyThrIleAsnGlyAsnGly
Db      607 GAGGACCTTAATATCTATCATCATGATGCGGCTCGCGGCTTGTATGACCAACGGA
Qy      179 GlnValTrpTrpProSerSerCysLysIleAsnLysSerLeuProCysArgAspIlePro
Db      667 AAAATCGGTGGCAAAATCATCATCAAAATCGAATATCAAGCATGCAAAAGCGCCA
Qy      199 ThrAlaLeuThrPheTrpPheCysLysAsnLeuLysValAsnAsnLeuLysSerLysAsn
Db      727 ACGGCTTACTCTCTACCAACTTAACAAATTTGAATGACAAATCTGAGAGTGAAGAAAT
Qy      219 AlaGlnGlnIleHisIleLysPheGluSerCysThrAsnValAlaLysSerLeuMet
Db      787 GCACACGACAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG
Qy      239 IleAsnAlaSerAlaLysSerProAsnThrAspGlyValHisValSerAsnThrGlnTyr
Db      847 ATCACTGCTCTCGGCATATGCTCCCAACAGATGATTCATCATCTGCTACTATAAAAC
Qy      259 IleGlnIleSerAspThrIleIleGlyThrGlyAspAspCysIleSerIleValSerGly
Db      907 ATTCGAATCTCAATTCAGACATTCGGACAGCTGATTCATTCATTCATTCATTCATTCATTC
Qy      279 SerGlnAsnValGlnAlaThrAsnIleThrCysGlyProGlyHisGlyIleSerIleGly

```

Db 967 TCGCAAAATGTTCAATCAATGATTAACTTGGGCCCCGCTCATGGCATCGACATGGA 1026
Qy 299 SerLeuGlySerGlyAenSerGluAlaTyrValSerAenValThrValangluAlaTyr 318
Db 1027 ACCTGGGGGAGATGACAAATTCCAAAGCTTATGTATCGGAATTAATGTGATGGTCTAG 1086
Qy 319 IleIleGlyAlaGluAenGlyValArgIleLeuThrTrpGlnGlyGlySerGlyGlnAla 338
Db 1087 CTCTGTGACACTGACAAATGAGTAAGAAATCAAGACTTACAGGAGAGGTCAAGAACTGCT 1146
Qy 339 SerAenIleLeuPheLeuAenValGluMetGlnAenValTyrProIleIleLeuAen 358
Db 1147 AAGAACTTAATTCAAATAATCTGTAATGATTAATGTAAGATTCGATTAATTCGAC 1206
Qy 359 GlnAenTyrCyAenAenGlyAlaGluProCyAlaGlnGlnPheSerAlaValGlnValTyr 378
Db 1207 CAGAACTACTGGCAAG---GACAAATGCGAAACAAGAAATCTCGGTTCAAGTGAAC 1263
Qy 379 AsnValValTyrGlnAenIleLeuGlyThrSerAlaThrLeuValAlaIleLeuPheAen 398
Db 1264 AATGTCGTATCGAAACATACAGGTACAGCGCAACGATGTGGCATTAATGTTTAAT 1323
Qy 399 CysSerThrAenPheProCysGluGlyIleIleMetGluAenIleAenLeuValGlyGlu 418
Db 1324 TGCAGTCAATATATCATGCCAGGTATGTGCTGAGAAATGTGAACATCAAGGA--- 1380
Qy 419 SerGlyLeuProSerGluAlaThrCyAlaAenValIlePheAenAenIleGluHisVal 438
Db 1381 -----GAAAGCTTCTTCAAAATGTCAATGTTAAGATTAAGGACCGTT 1428
Qy 439 ThrProHisCys 442
Db 1429 TCTCTTAATGC 1440
RESULT 9
US-10-437-963-24607
; Sequence 24607, Application US/10437963
; Publication NO. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 24607
; LENGTH: 1182
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_29575C.1
US-10-437-963-24607
Alignment Scores:
Pred. No.: 5,55e-107 Length: 1182
Score: 1011.00 Matches: 187
Percent Similarity: 69.52% Conservative: 57
Best Local Similarity: 53.28% Mismatches: 105
Query Match: 42,304 Indels: 2
DB: 18 Gaps: 1
US-10-691-374-2 (1-457) x US-10-437-963-24607 (1-1182)
Qy 72 GlyIleLeuValIleAenValLeuSerPheGlyAlaLeuGlyAenGlyLeuThrTyrAenP 91

Db 79 GGCAGTAAATGTTTACGACATACAGAGCTACGGGGCTCATGAGAGAGGACATGATGAC 138
Qy 92 AenIleAlaPheGluGlnAlaTyrAenGluAlaCysSerSerAenThrProValGlnPhe 111
Db 139 ACCAAAGCATTTGGAGATACATGAGCTGACGCTGCTCTCTGCAAAACCTTCAGATTTCG 198
Qy 112 ValValProLeuAenIleAenThrLeuLeuValGlnIleThrPheSerGlyProCysAen 131
Db 199 CTCATCCCAAGAGGAGAAATACCTGATCAAGACACAAACCTGCTGATCCATGACAA 258
Qy 132 SerSerIleSerValIlePheGlySerLeuGluAlaSerSerIleSerAenTyr 151
Db 259 TCATACATCTCATTTGATGTGAAGGTATGTTGATGGCTTCCAGAGAGGTCAAGCTGG 318
Qy 152 LysAenAenG---ArgLeuThrIleAlaPheAenSerValGlnAenLeuValGly 169
Db 319 ACCAAGAGACCATTAAGGACATGATCTGATCACTGTGTCTCACTGTCTCTCACT 378
Qy 170 GlyIleGlyThrIleAenGlyAenGlyValGlnValThrProSerSerCysAenIleAen 189
Db 379 GGTGTGGGACCATTAAGTAAGAAATGCAAGATTTGATGGCAAAATTCATGCAAAACCAAC 438
Qy 190 LysSerLeuProCysAenAenAenAenAenAenAenAenAenAenAenAenAenAenAen 209
Db 439 TCAAGCTTCCATGACACAGAGCTCAAGCGCATGATCTACTCTGCAAGAAATCTG 498
Qy 210 LysValAenAenLeuLeuSerLeuAenAenAenAenAenAenAenAenAenAenAenAen 229
Db 499 AAGTAGAGATATCTGAAGGTGTGAAACGCCAATTCAGATTTCAAGTGAAGATTCG 558
Qy 230 ThrAenValValAlaSerAenLeuMetIleAenAlaSerAlaLeuSerProAenThrAen 249
Db 559 ACCGATGATGATGCTCTGCTGCTGATCACAGCAACCAAGACCCCAACACTGAT 618
Qy 250 GlyValHisValSerAenThrGlnTyrIleGlnIleSerAenThrIleGlyThrGly 269
Db 619 GGAATTCATATCACACGACAGAGATGTGAGGTGACAGATGATGATCAAGACCGG 678
Qy 270 AenAenCysIleSerIleValSerGlySerGlnAenValGlnAlaThrAenIleThrCys 289
Db 679 GATGACTCATGTCAATCGAGGACGAAACCGAATCTCATGTCAAGAACTGTGTGT 738
Qy 290 GlyProGlyHisGlyIleSerIleGlySerLeuGlySerGlyAenSerGluAlaTyrVal 309
Db 739 GACCCGGGACAGGATCAGATCGATCGGTGCTGGGTGATCATTAATTCGAACTCATGTC 798
Qy 310 SerAenValThrValAenGluAlaLeuValIleIleGlyAlaGluAenGlyValArgIleLeu 329
Db 799 AACATGTACCTGACATGATCATGCTGATGACAGCCAAACGAGACTGCAATCAAG 858
Qy 330 ThrTrpGlnGlyIleSerGlyGlnAlaSerAenIleLeuPheLeuAenValGluMetGln 349
Db 859 ACATGCGAGGAGAAAGGTTCAAGCAAAATCTGTTCCAGAACTGTCTCATGTGAC 918
Qy 350 AsnValLeuTyrProIleIleIlePheGlnAenTyrCyAenAenValGluProCysIle 369
Db 919 AATGTCGAAACCCATCATCATCAACAAACTACTGAGCTTCTTAACCCCTGCAAG 978
Qy 370 GlnGlnPheSerAlaValGlnValIleAenValIleTyrGluAenIleLeuGlyThrSer 389
Db 979 CAACAGAAATCTGACGTGAGAGGTGACATGCTGCTTCAAAATCAACGAGGACCAAT 1038
Qy 390 AlaThrLysValAlaIleLeuPheAenCysSerThrAenPheProCysGluGlyIleIle 409
Db 1039 GCATAGAGAGAGGACATCTGCTGATTCAGACAGCAAGTATCTTGACATGCAATAAC 1098
Qy 410 MetGluAenIleAenLeuValGlyGluSerGly 420
Db 1099 TTGAGAAATGTCAATCTCACTGTCAAGGAGGA 1131
RESULT 10
US-10-437-963-24608

```
/ Sequence 24608, Application US/10437963
/ Publication No. US20040123343A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovalic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Wu, Wei
/ APPLICANT: Boukharov, Andrey A.
/ APPLICANT: Barbazuk, Brad
/ APPLICANT: Li, Ping
/ TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(53221)B
/ CURRENT APPLICATION NUMBER: US/10/437,963
/ NUMBER OF SEQ ID NOS: 204966
/ SEQ ID NO 24608
/ LENGTH: 1359
/ TYPE: DNA
/ ORGANISM: Oryza sativa
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MKT4530_29576C.1
US-10-437-963-24608
```

```
Alignment Scores:
Pred. No.: 9.14e-107 Length: 1359
Score: 1010.00 Matches: 190
Percent Similarity: 70.11% Conservative: 54
Best Local Similarity: 54.60% Mismatches: 102
Query Match: 42.26% Indels: 2
DB: 18 Gaps: 1
```

US-10-691-374-2 (1-457) x US-10-437-963-24608 (1-1359)

```
QY 75 ValIleAsnValLeuSerPheGlyAlaLysGlyAspGlyThrTrpAspAsnIleAla 94
DB 154 GTGTCAGATGTCGCAAAACATGCGCTTACGAGATGCAACATGATGACAGAAAGCA 213
QY 95 PheGlnGlnAlaTrpAsnGlnAlaCysSerSerArgThrProValGlnPheValAlaPro 114
DB 214 TTGGCAAAAGCATGGCGTGCAGCTGCTCTTCGCAACCTTCATGTTGTCATCCCA 273
QY 115 LysAsnLysAsnTrpLeuLeuLysGlnIleThrPheSerGlyProCysArgSerSerIle 134
DB 274 AAGGGCAAGAGATACCTCAACCAAGCATATACATCTGTCATGATGCAAAATCAAGCATC 333
QY 135 SerValLysIlePheGlySerLeuGlnAlaSerSerLysIleSerAspTrpLysAspArg 154
DB 334 ACCTTCATGATAGAGGTACGTTGCGCTCCTCCAAAGAGGTCAAGATTGGAGCAAGCA 393
QY 155 -----ArgLeuTrpIleAlaPheAspSerValGlnAsnLeuValAlaGlyGly 172
DB 394 ACTATTAGGACGACGATATGTTCAATGATGATGAGTGAAGGCTTACGTCCTGCTGATGG 453
QY 173 ThrIleAsnGlyAsnGlyGlnValIleTrpProSerSerCysLysIleAsnLysSerLeu 192
DB 454 ACTGTCGATGAAGAAACGCAAGATTGGTGCAAATTCCTGCAAAACGAATGCAGAACTT 513
QY 193 ProCysArgAspAlaProThrAlaLeuThrPheTrpAsnCysLysAsnLeuValAlaAsn 212
DB 514 CCATGACATGAAGCTCCAAAGCGCTTGAATTCCTGTTCAAACTGAAAGTGAG 573
QY 213 AsnLeuLysSerLysAsnAlaGlnGlnIleHisIleLysPheGlnLysCysThrAsnVal 232
DB 574 AACTTGAACTACTAAACAGCCAAACAAATCCCATGATGATGAGATTGACATGATGTT 633
QY 233 ValAlaSerAsnLeuMetIleAsnAlaSerAlaLysSerProAsnThrAspGlyValHis 252
DB 634 AGCATCTAGCCATCAATCAACGACCAAGGACCTACGCAAAACGAGGAGGATCAT 693
QY 253 ValSerAsnThrGlnTrpIleGlnIleSerAspThrIleIleGlyThrGlyAspAspCys 272
DB 272 ValSerAsnThrGlnTrpIleGlnIleSerAspThrIleIleGlyThrGlyAspAspCys 272
```

```
DB 694 ATCAAGAGTAAATATGTACAGTACAGCGCTGCATATATCAAGCCGGGATGACTGC 753
QY 273 IleSerIleValSerGlySerGlnAsnValGlnAlaThrAsnIleThrCysGlyProGly 292
DB 754 ATGTCATGATGAGGATGAACTGGAACCTTACATGTCAAGACATGTTGTGTGACCGGGA 813
QY 293 HisGlyIleSerIleGlySerLeuGlySerGlyAsnSerGlnAlaTrpValSerAsnVal 312
DB 814 CACGCGATACGATCCGGAGCTTAGCGCATCAACACTGAAAGCTCATGATCAACATGTTC 873
QY 313 ThrValAsnGlnAlaLysIleIleGlyAlaGlnAsnValAlaArgIleLysThrTrpGln 332
DB 874 ACCGTGACACCGTCAGGCTATATGACCAACCAAGGAGCTCCGATCAAGACATGGAG 933
QY 333 GlyGlySerGlyGlnAlaSerAsnIleLysPheLeuAsnValGlnMetGlnAspValLys 352
DB 934 GGGGTTGGGGCTTACCGCAAGAACATGCTGTCCGAACATGATCATGAAAAATGTTGG 993
QY 353 TyrProIleIleIleAspGlnAsnTrpCysAspArgValGluProCysIleGlnGlnPhe 372
DB 994 AACCCATCATATATGACCAAGACTACTGTGATTTGCTACACCTTGCAAGAAACAGCA 1053
QY 373 SerAlaValGlnValLysAsnValAlaTrpGlnAsnIleLysGlyThrSerAlaThrLys 392
DB 1054 TCTGCGGTGCAAGTTAGCAATGTGTCTTCAAGAACATCAGGGGCAACAGTCTTCCAG 1113
QY 393 ValAlaIleLysPheAspCysSerThrAsnPheProCysGlyGlyIleIleMetGlnAsn 412
DB 1114 GAGGCTATCAAAACGAGCTGATGACGAATATCTTGTCAAGAAATTAACCTTGAAAGAC 1173
QY 413 IleAsnLeuValGlyLysGly 420
DB 1174 GTCAAGCTCACTATCAAGGAGGT 1197
```

RESULT 11

US-10-437-963-28561

/ Sequence 28561, Application US/10437963

/ Publication No. US20040123343A1

/ GENERAL INFORMATION:

/ APPLICANT: La Rosa, Thomas J.

/ APPLICANT: Kovalic, David K.

/ APPLICANT: Zhou, Yihua

/ APPLICANT: Cao, Yongwei

/ APPLICANT: Wu, Wei

/ APPLICANT: Boukharov, Andrey A.

/ APPLICANT: Barbazuk, Brad

/ APPLICANT: Li, Ping

/ TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

/ FILE REFERENCE: 38-21(53221)B

/ CURRENT APPLICATION NUMBER: US/10/437,963

/ NUMBER OF SEQ ID NOS: 204966

/ SEQ ID NO 28561

/ LENGTH: 1335

/ TYPE: DNA

/ ORGANISM: Oryza sativa

/ FEATURE:

/ OTHER INFORMATION: Clone ID: PAT_MKT4530_33148C.1

US-10-437-963-28561

```
Alignment Scores:
Pred. No.: 2.84e-104 Length: 1335
Score: 988.50 Matches: 186
Percent Similarity: 68.88% Conservative: 73
Best Local Similarity: 49.47% Mismatches: 110
Query Match: 41.36% Indels: 7
DB: 18 Gaps: 5
```

US-10-691-374-2 (1-457) x US-10-437-963-28561 (1-1335)

```
QY 75 ValIleAsnValLeuSerPheGlyAlaLysGlyAspGlyThrTrpAspAsnIleAla 94
DB 94 ValIleAsnValLeuSerPheGlyAlaLysGlyAspGlyThrTrpAspAsnIleAla 94
```



```
Db      155 GAGAGGACGAGCTACCTGCTCCGCAAGTACACCTTCCGCGCCCTGCAATCCACATC 214
Qy      135 SerVallylIlePheGlySerLeuGlnAlaSerSerIleIleSerAspTyrIleAsp--- 153
Db      215 AACCTAGTGTGAAGGGAGCGCTGGTGGCTCCGCGCATGTGCAACTGCAACGAGAGC 274
Qy      154 ---ArgArgLeuTrpIleAlaPheAspSerValGlnAsnLeuValGlyGlyGly 172
Db      275 AACAGAGGTACTGATCTCTCCGCGCGGTGAGAGACTCTCGCAAGATCAACGGGCATC 334
Qy      173 ThrIleAsnGlyAsnGlyGlnValTyrTrpProSerSerCysIleIleAsnIleSerLeu 192
Db      335 ACCATGACGCGCAACGCGGAGGTGGTGGAGAACTCTCGCAAGATCAACGGGCATC 394
Qy      193 ProCysArgAspAlaProThrAlaLeuThrPheTrpAsnCysIleAsnLeuValAsn 212
Db      395 CCATGCAAGGAGCTCCACGCGCTGAGTTTCCACAGTGGACACCTGAGCGGTGAAT 454
Qy      213 AsnLeuIleSerIleAsnAlaGlnGlnIleHisIleIlePheGlnSerCysThrAsnVal 232
Db      455 GGTCTGAGATGATGTAACAGCAGAGATCCACATGTAGTGGAGATTCACCGGGGTG 514
Qy      223 ValAlaSerAsnLeuMetIleAsnAlaSerAlaIleAsnIleThrAsnIleThrCysGlyProGly 252
Db      515 GAGCTGGCCACCTGTCAATCTGTGCGCCGCGCAAGAGCCCAACAGATGGCATCAC 574
Qy      253 ValSerAsnThrGlnTyrIleGlnIleSerAspThrIleIleGlyThrGlyAspAspCys 272
Db      575 ATCAACCAACAGCAAGAGATGCAAGTCCAGATGCACTGCACTTAAGACAGGGGATGACTGT 634
Qy      273 IleBerIleValSerGlySerGlnAsnValGlnAlaThrAsnIleThrCysGlyProGly 292
Db      635 GGTGTGATGAGAGATGGAGCCACCGCTTTCACCTCCAGACTGCTGTGTGGGCCCGGG 694
Qy      293 HisGlyIleSerIleGlySerLeuGlySerGlyAsnSerGlnAlaTyrValSerAsnVal 312
Db      695 CATGGAGATTCATTTGGAGCTTGGAGATGACAACTCCAGAGTGGAGTGTGACATC 754
Qy      313 ThrValAsnGlnAlaIleIleIleGlyAlaGlnAsnGlyValArgIleIleThrTPGln 332
Db      755 TTCATGCAACCGTGACCTCTATGAGCACCACCAATGAGACTGGATCAAGATGCGAG 814
Qy      333 GlyIleSerGlyGlnAlaSerAsnIleIlePheLeuAsnValGlnMetGlnAspValIle 352
Db      815 GAGGAGATGATTCGCAAGATATGTATTCAGAAACATGGTCAACAGTGTCAAG 874
Qy      353 TyrProIleIleIleAspGlnAsnTyrCysAspArgValGlnProCys---IleGlnGln 371
Db      875 AACCCATAATCATTTGACCAAACTACTGCACTCAGCTTAAGAAATGCCAGACAGCGAG 934
Qy      372 PheSerAlaValGlnValIleAsnValAlaTyrGlnAsnIleIleGlyThrSerAlaThr 391
Db      935 GGTACGACATGAGATTCAGAAATGTGCTTTCACAAACATCGCAGGGGCAACAATTTCC 994
Qy      392 LysValAlaIleIlePheAspCysSerThrAsnPheProCysGlnGlyIleIleMetGln 411
Db      995 AAGAGTGCCTACTCTGAACTGCAAGACATACCTCATGTATGACATTTCTTACG 1054
Qy      412 AsnIleAsnLeu-----ValGlyIleSerGlyIlePheSerGlnAlaThrCysIleAsn 429
Db      1055 GACATCAACCTGAAATGTGTGATGCAATGTGT---GCCACGAAAGATCTTGCAGAAAT 1111
Qy      430 ValHisPheAsnAsnAlaGlnIleIleValThrProHis---CysThrSer 444
Db      1112 GCMAAATGAGAAATCTGAAACATGTTTCCACACCAATGACACGAGC 1159

RESULT 13
US-10-425-115-33752
; Sequence 33752, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
```

```
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21 (53222) B
; CURRENT APPLICATION NUMBER: US/10/425,115
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 33752
; LENGTH: 1372
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_130788C.1
US-10-425-115-33752

Alignment Scores:
Pred. No.: 1,33e-98 Length: 1372
Score: 940.00 Matches: 174
Percent Similarity: 71.22% Conservative: 66
Best Local Similarity: 51.63% Mismatches: 93
Query Match: 39.33% Indels: 4
DB: 18 Gaps: 3

US-10-691-374-2 (1-457) x US-10-425-115-33752 (1-1372)

Qy      114 ProLysAsnIleAsnTyrIleuLeuIleGlnIleThrPheSerGlyProCysArgSerSer 133
Db      12 CTGCGCGGAGCGGTACTCTCTGAGCTCTACACCTCCGTGGCCGTGCAAGTCCAGC 71
Qy      134 IleSerValIlePheGlySerLeuGlnAlaSerSerIleSerAspTyr----- 151
Db      72 GTACAGCTCAGCGAGAAAGGACCCCTGTGCTGATGCAACCGCGGAGTTGACCGAC 131
Qy      152 LysAspArgLeuTrpIleAlaPheAspSerValGlnAsnLeuValIleGlyGly 171
Db      132 AATGACAGAGGACACTGATGTGTTCCGAGCATGCAACAGCTCACTGCAACGGCGGT 191
Qy      172 GlyThrIleAsnGlyAsnGlyGlnValTyrTrpProSerSerCysIleAsnIleSer 191
Db      192 GCGCCATCGATGAGCAACGCGCGAAGATGTGGCTTCACTCGTCAAGATCAACAGCT 251
Qy      192 LeuProCysArgAspAlaProThrAlaLeuThrPheTrpAsnCysIleAsnLeuVal 211
Db      252 CTCCCTTGCAGAGAGCTTCGACGCGCTGTCTCATTCACCTGCTGCGTCAACCTGAAGGTG 311
Qy      212 AsnAsnLeuIleSerIleAsnAlaGlnGlnIleHisIleIlePheGlnSerCysThrAsn 231
Db      312 GAGAACTGAAATTTATGAACAGCAACAGATCAATGTCACTGACAGACTGTGCMAAT 371
Qy      232 ValIleAsnSerAsnLeuMetIleAsnAlaSerAlaIlePheSerProAsnThrAspGlyVal 251
Db      372 GTGCTGTGCGCCGGTGTTCATCACAGCGCTGACACCGCTTAACCTGACCGCATC 431
Qy      252 HisValSerAsnThrGlnTyrIleGlnIleSerAspThrIleIleGlyThrGlyAspAsp 271
Db      432 CACATCACTCGCAACAGATGATGCGCTCAACAGCTGCAAGATCAAGACAGAGAGCAGC 491
Qy      272 CysIleSerIleValSerGlySerGlnAsnValGlnAlaThrAsnIleThrCysGlyPro 291
Db      492 TGCATGTCAATCAGAAACGGAAGCTCAACACTCTCATGTCTCAAGATTACTGTGTCA 551
Qy      292 GlyHisGlyIleSerIleGlySerLeuGlySerGlyAsnSerGlnAlaTyrValSerAsn 311
Db      552 GGGCATGGGATGATGATCGAGATCGCTTAGAGACGACAACTCAAGAGCGAAGTCTCCGCG 611
Qy      312 ValThrValAsnGlnAlaIleIleIleGlyAlaGlnAsnGlyValArgIleIleThrTrp 331
Db      612 ATCAACCATGATTCATGTCAGCACTCAACGCAACCAACGAGGACCGCATCAAGAGTAC 671
Qy      332 GlnGlyIleSerGlyGlnAlaSerAsnIleIlePheLeuAsnValGlnMetGlnAspVal 351
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OW protein - nucleic search, using frame_plus_p2n model

Run on: March 4, 2005, 12:56:14 ; Search time 4970 Seconds
(without alignment)

4455.540 Million cell updates/sec

Title: US-10-691-374-2

Sequence: 1 MVIQRNLSLLILITIPASSIS.....VTPHCTSLIESEDEALVNY 457

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODL=frame+ p2n.model -DEV=xlp
-Q/cgn2.1/USPTO spool.p/US10691374/runat 28022005 120705 20986/app_query.fasta_1.647
-DB=genemb1 -QFWT=fastcap -SUFFIX=rgc -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human4.cdf -LIST=45
-DOCALLIGN=200 -THR.SCORE=PCT -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10691374 @CGN 1.1 5600 @runat 28022005 120705 20986 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -DSPBLOCK=100 -LONGCLOC
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmb1:*
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pal:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_pro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	Score	Query	Match Length	ID	Description
No.					
1	2390	100.0	1617	8	LEPER
2	2390	100.0	1621	6	AX05656 Tomato mRNA
3	2390	100.0	1621	6	AX15981 L. esculentu
4	2390	100.0	1621	8	AX062336 Sequence
					LEPG2AR
					AX0583 Tomato mRNA

5	2390	100.0	1624	6	I01809
6	2390	100.0	1636	6	A24194
7	2390	100.0	1636	6	AB364905
8	1506	63.0	7456	8	LEPOLYGA
9	1506	63.0	7456	8	TOMPGANA
10	1282	53.6	1541	8	AF152758
11	1204.5	50.4	1595	8	AY043233
12	1153	48.2	1725	8	AY06POLYGA
13	1149	48.1	1364	8	AY078936
14	1136	47.5	1755	8	PAPOLYGA
15	1130	47.3	1765	8	BNA250919
16	1128	47.2	1767	8	AF062467
17	1127	47.2	1296	6	AX412563
18	1127	47.2	1296	6	AX651724
19	1127	47.2	1296	6	BT005376
20	1127	47.2	1296	6	AK117942
21	1115.5	46.7	1280	6	AX651722
22	1113	46.6	1599	8	BMPGALACR
23	1113	46.6	1631	6	AE2402
24	1106	46.3	1302	8	BPA428543
25	1103	46.2	1657	6	AS7101
26	1103	46.2	1657	6	AX016328
27	1103	46.2	1657	6	BNSAC66
28	1092	45.7	1773	8	MAUP14A
29	1074	44.9	1795	8	AB084461
30	1061.5	44.4	1208	8	R1MJ4147
31	1059	44.3	599	8	CPA505750
32	1044	43.7	1641	8	PC0504855
33	1022.5	42.8	1269	6	AX653992
34	1011	42.3	1182	6	AX653993
35	1005	42.1	602	8	LES505947
36	999	41.8	1344	6	AX654086
37	925	38.7	1603	8	AF138858
38	914	38.2	1366	8	AY142668
39	914	38.2	1471	8	AY046002
40	911	38.1	1450	8	AK105520
41	868.5	36.3	87792	8	AP003837
42	852.5	35.7	4440	8	PPPG
43	849	35.5	87792	8	AP003837
44	849	35.5	120823	8	AP003748
45	841.5	35.2	1236	6	AX652801

ALIGNMENTS

RESULT 1	LEPER	1617 bp	mRNA	linear	PLN 12-SEP-1993
LOCUS	Tomato mRNA for polygalacturonase (PG, EC 3.2.1.15).				
DERIVATION	LEPER				
ACCESSION	X05656				
VERSION	X05656.1	GI:19297			
KEYWORDS	polygalacturonase.				
SOURCE	Lycopersicon esculentum	(tomato)			
ORGANISM	Lycopersicon esculentum				
REFERENCE	Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
AUTHORS	Sheehy,R.R., Pearson,J., Brady,C.J. and Hiatt,W.R.				
TITLE	1 (bases 1 to 1617)				
REFERENCE	2 Rogers,H.J., Allen,R.L., Hamilton,W.D. and Lonsdale,D.M.				
AUTHORS	Pollen specific cDNA clones from Zea mays				
TITLE	Biotechm. Biophys. Acta 1089 (3), 411-413 (1991)				
JOURNAL	91316147				
MEDLINE	1859845				
PUBMED	3 (bases 1438 to 1438)				
REFERENCE	Hiatt,W.R.				
AUTHORS	Submitted (06-OCT-1987)				
TITLE	Location/Qualifiers				
JOURNAL	1. 1617				
FEATURES	source				

```

/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/db_xref="taxon:4081"
/clone="F1"
47..1420
/feature="unlabeled protein product; polygalacturonase (AA
1-457)"
/codon_start=1
/protein_id="CAA29148.1"
/db_xref="GI:19298"
/db_xref="GOA:P05117"
/db_xref="UniProt/Swiss-Prot:P05117"
/translation="MVIQNSILLIITIPASSISTGCSNYIDNLFKQYVNTLLEQF
ADPQAVLSYLSKIESSNNIDKVDKGIIVINLSPGAKGDKTYDNLAREANMEA
CSRTVPQVFPVPRKNVYLKQIIPSGGCRSISVYKIFGSLBASSTKSDTKDRMLAF
DSYQNLVYGGGRTINGQYWPSPSSCKINKSLPORDAPLALITWNCNKLVNLSKN
AQQHLIFRSECTNVASNLININASAPNTDGVASVSTQYIQISDPIIGGDDCISIV
SGSONVATATITGPGHGISIGLSGNSSEAVSNVYNAKTIQAGENVRIKTMQGG
SGQASNIKPLNVEDVKPPIIDQNYCDEBPCIOQFSAVQVNVVENEIKGTSATK
VAIKFDCSTNPFCEGIIMENINLVSGSGKSEATCKNVHNNAEHVTPHCTSLSEISD
EALLVNY"
1438
/feature="C was g in [1]"
/citation=[1]
1465..1470
/feature="pot.polya signal"
1597..1603
/feature="pot.polya signal"

ORIGIN

Alignment Scores:
Pred. No.: 6.24e-192 Length: 1617
Score: 2390.00 Matches: 457
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 8 Gaps: 0

US-10-691-374-2 (1-457) x LEPGR (1-1617)

QY 1 MetValIIeGIArGaSenSerIIeLeuLeuLeIIeIIePheAlaSerSerIIeSer 20
Db 47 ATGGTTATCCAAAGAAATGATTTCTCTTCATATTAATTTGCTTCATCAATTTCA 106
QY 21 ThrCyAsrSerAsnValIIeAspAspAsnLeuPheGlnValIlyrAspAsnIIeLeu 40
Db 107 ACTGTAGAAGCAATGTATTGATGACATTTATTCAACAAGATTATGATATAATTTCT 166
QY 41 GlnGlnGlnPheAlaHISpPheGlnAlaIlyrLeuSerIlyrLeuSerIlyAsnIIeGln 60
Db 167 GAACAAGAAATTTGCTCATATTTTCAAGCTTATCTTTTATTGAGCAAAAATATTGAA 226
QY 61 SerAsnAsnAsnIIeAspIlyValAspIlyAsnGlyIlyLeuValIIeAsnValLeuSer 80
Db 227 AACCAACAATTAATTTGACAAAGTTGATATAAATGGCATTAAGTATTAATGACTTAC 286
QY 81 PheGlyAlaIlyGlyAspGlyIlyrThrIlyrAspAsnIIeAlaPheGlnGlnAlaIlyrAsn 100
Db 287 TTTGAGGCTAAGGGGATGAGAAACATATGATATATTGACTTTGAGCAAGCATGAGAT 346
QY 101 GlnAlaCySerSerSerIlyrProValGlnPheValIlyrProValAsnIlyAsnIlyLeu 120
Db 347 GAGCATGTTCACTTAGAACCTCTTTCATTAATTTGGTCTCTTAACAAACAAATTAATCT 406
QY 121 LeuIlyGlnIIeThrPheSerGlyProCyAsrSerSerIIeSerValIlyeIIePheGly 140
Db 407 CTCAGCAAAATCACTTTTCAAGTTCATGACGATTTCTATTTCAGTAAAGATTTTGGGA 466
QY 141 SerLeuGlnIlyAsrSerIlyeIIeSerAspIlyrIlyAspAspAqArGLeuTrpIIeAlaPhe 160
Db 467 TCTTGAAGACATCTAGTAAATTTTCAACATTAAGATAGAGGCTTTGGATGCTTTT 526
QY 161 AspSerValGlnAsnLeuValIlyGlyIlyGlyIlyThrIIeAsnGlyAsnGlyGlnVal 180

```

```

Db 527 GATAGTGTCAAATTAAGTTGTTGGAGAGAGGAACTATCAATGGCAATGACAAAGTA 586
QY 181 TrpIlyrProSerSerCyIlyrIlyeAsnIlySerIlyeProCyAsrAspAlaProThrAla 200
Db 587 TGGTGGCCAAAGTTCTTGCAAAATTAATTAATCACTGACCATGAGGAGGACCAACGGCC 646
QY 201 LeuThrPheTrpAsnCyIlyrAsnLeuIlyValAsnAsnLeuIlySerIlyAsnIIeGln 220
Db 647 TTAACCTTCTGGAATTCGCAAAAATTTGAAGGATATATCTTAAGAGTAAATATCACAA 706
QY 221 GlnIIeHISIIeIlyrPheGlnIlySerCyThrAsnValIlyAlaSerAsnLeuMetIIeAsn 240
Db 707 CAAATTCATATCAAAATTTGAGTCAGCACTATATGTTGAGCTTCAAAATTTGATGATCAT 766
QY 241 AlaSerAlaIlySerProAsnThrAspGlyValHISerAsnThrGlnTrpIlyeGln 260
Db 767 GCTTCAGCAAGAGCCCAATTCATGATGAGATCCATGATCAATATCTCAATATATTCA 826
QY 261 IlySerAspThrIIeIlyGlyThrGlyAspAspCyIlySerIlyeAlaIlySerGln 280
Db 827 ATATCTGATTAATTAATTTGAAACAGTGAATGATTTGATTTCAATGTTTCTGATCTCAA 886
QY 281 AsnValGlnAlaIlyrAsnIIeThrCyGlyIlyrProGlyHISerIlyeGlySerIlyeLeu 300
Db 887 AATGTGCAAGGCACAAATTAATTAATCTTGCTGCTCAGTCAATGATTAAGTATTTGAAAGCTTA 946
QY 301 GlySerGlyAsnSerGlnIlyrValIlySerAsnValIlyrValAsnGlnAlaIlyeIlye 320
Db 947 GATCTGGAATTTCAAGAAAGCTTAATGATGATGATTTATCTGATTAATGAAGCCAAATTAATTC 1006
QY 321 GlnValGlnGlnGlyValIlyGlyIlyeThrTrpGlnGlyIlySerGlyGlnAlaSerAsn 340
Db 1007 GGTGCGCAAAATTTGAGTATGATCAACACTTGCGCAGGAGATCTGCAACAGTACGCAAC 1066
QY 341 IlyeIlyrPheLeuAsnValGlnMetGlnAspValIlyrTrpIlyeIlyeIlyeAspGlnAsn 360
Db 1067 ATCAATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1126
QY 361 TrpCyAspArgValGlnProCyIlyeGlnGlnPheSerAlaValGlnValIlyrAsnVal 380
Db 1127 TATGTGATGAGTGAACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1186
QY 381 ValIlyrGlnAsnIlyeGlyIlyrSerAlaIlyrIlyrValAlaIlyrPheAspCySer 400
Db 1187 GTGATAGAAATATCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAG 1246
QY 401 ThrAsnPheProCyGlnGlyIlyeIlyeMetGlnAsnIIeAsnLeuValIlyGlyIlySerGly 420
Db 1247 ACAACTTTCATGATGAGAAATTAATTAATGAGAAATTAATTAATTAATTAATTAATTAATTA 1306
QY 421 IlyrProSerGlnAlaThrCyIlyrAsnValHISpAsnAsnAlaGlnHISValIlyrPro 440
Db 1307 AAACCATCAAGAGCTACGTGCAAAATTTGCTCAATTTTAAACAATGCTCAACAATGTTTACACA 1366
QY 441 HISerThrSerLeuGlnIlyeSerGlnAspGlnAlaIlyeLeuIlyrTrpAsnIlyr 457
Db 1367 CACTGCACTTCACTAGAAATTTCAAGAGATGAAGCTCTTTGTATATTAATTA 1417

RESULT 2
A15981 1621 bp RNA linear PART 02-MAR-1994
LOCUS L.esculentum mRNA for polygalacturonase-2a.
DEFINITION A15981
ACCESSION A15981
VERSION A15981.1 GI:490039
KEYWORDS
SOURCE
ORGANISM
Lycopersicon esculentum (tomato)
Lycopersicon esculentum
Rutaceae; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 1621)
REFERENCE
Bridges,I.G., Schuch,W.W. and Grierson,D.

```

TITLE Anti-sense regulation of plant gene expression
JOURNAL Patent: EP 0271998-A 1 22-JUN-1986;
IMPERIAL CHEMICAL INDUSTRIES PLC
FEATURES
source

1..1621
/organism="Lycopersicon esculentum"
/mol_type="unassigned RNA"
/strain="Alisa Craig"
/db_xref="taxon:4081"
/clone="PTOM 6"
/cissue_type="fruit"
/dev_stage="fruit ripe"
51..1424
/codon_start=1
/product="polygalacturonase-2a"
/protein_id="CAA01256.1"
/db_xref="GI:490040"
/db_xref="GOA:P05117"
/db_xref="UniProt/Swiss-Prot:P05117"
/translation="NVIQRNSILLITIPASSISTCRNVIDDNLFKOYDNIIEQEF
AHDFOAYLSYLSKNISNNNDKVDKNGIKYINVLFSKAGDKTYDNIADFOANVEA
CSRPVQVVPVVKNNKYLKQTFSGPKRSISVRI FGLSESSKISDYKORLWIAF
DSVQLVVGGGTNGVQWVPSCKINKSLPCRDAPLALFNNCKMLKYNNLSKN
AQQHFKESCTNVAVASMTNASKSPNTDGVHNTQYIOISPTIIGTGDICISIV
SGSNVQNTNITCGPHGISIGSGNSEAVSVTVTEAKIIGAENVRIKTMQGG
SGQSNITKFLAVEMQDVKPIITIDNTCDRPEPCIQPSAVQKRVYENIKGYSATK
VAIKFDCSTNPFCEGIIMENINLVGSKSPSEATCKNVPNNAEHVTHCSLEISD
EALVNYN"

mat_peptide
264..1421
/product="polygalacturonase-2a"
615..623
/note="pot. N-glycosylation site"
misc_feature
768..776
/note="pot. N-glycosylation site"
misc_feature
906..914
/note="pot. N-glycosylation site"
misc_feature
981..989
/note="pot. N-glycosylation site"
misc_feature
1598..1606
/note="pot. N-glycosylation site"
misc_feature
1621
/note="pot. polyA signal"
polyA_site
1621
/note="polyA site"

ORIGIN

Alignment Scores:

Pred. No.: 6,266-192 Length: 1621
Score: 2390.00 Matches: 457
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-691-374-2 (1-457) x A15981 (1-1621)

QY 1 MetValIIEGlnARgsnSerIleleuLeuIleIleIlePheAlaSerSerIleSer 20
DB 51 ATGTTATTCCTCAAGAAATGATATTCCTTCCTCATTAATATTTGCTTCATCAATTTCA 110
QY 21 ThrCyARgsrAsnValIleAspAsnLeuPheValIleValIleAspAsnIleleu 40
DB 111 ACTTTAGAGCAATGTTATTTGATGCAATTTATTCBAACAAGTTATATGATAATTTCTT 170
QY 41 GIUGInGluPheAlaHleAspPheGlnAlaIleIleuSerIleuSerIleuSerIleu 60
DB 171 GAACAAGAAATTTGCTCATGATTTTTCACGCTTATCTTTCTTATTTGACAAAAATATTGA 230
QY 61 SerAsnAsnAsnIleAspIleValIleAspIleValIleValIleAsnValIleuSer 80
DB 221 AGCAACAATTAATATTGACAAGGTGATTAATAATGGATTAAGTATTAATGACTTAC 290
QY 81 PheGluAlaIleValIleAspGluIleValIleValIleAspAsnIleAlaPheGluGlnAlaTrpAsn 100
DB 221 TTGGAGCTAAGGATGATGGAATAACATATGATTAATTTGATTTGACACAAGCATGGAAT 350

QY 101 GIuAlaCySerSerArgThrProValGlnPheValValProIleAsnIleAsnIleu 120
DB 351 GAAGCATTTTCATCTAGAACACCTGTTCAATTTGGTCTCTAAATAACAGAAATATCTT 410
QY 121 LeuIleValGlnIleThrPheSerGluProCyARgsrSerIleSerValIleIlePheGlu 140
DB 411 CTCAGCAAAATCACTTTTCAGTCCATGCAAGATCTTCTATTTCAAGTAAAGTTTGTGA 470
QY 141 SerLeuGluAlaSerSerIleIleSerArgIleValAspArgArgLeuTrpIleAlaPhe 160
DB 471 TCCTTAAGAAGCATCTAGTAATTTCAAGATCAAAAGTAGAAGGCTTTGATTCCTTTT 530
QY 161 AspSerValGlnAsnLeuValValGluValGluValIleValIleAsnGluValGlnVal 180
DB 531 GATAGTGTTCAAAATTTAGTTGTGAGAGAGAGAACTATCAATGGCAATGACAGTA 590
QY 181 TrpTrpProSerSerCysIleValIleAsnIleuProCyARgsrAspAlaProThrAla 200
DB 591 TGGTGGCCAAAGTTCTTGCAAAATTAATTAATCATCGCATGCAAGGATGCACCAAGGCC 650
QY 201 LeuThrPheTrpAsnCysIleValAsnLeuIleValAsnLeuIleuSerIleAsnIleGln 220
DB 651 TTAACCTTCTGGAAATTCGAAAAATTTGAAGTGAATTAATCTAAAGATTAATAATGACACA 710
QY 221 GlnIleHleIleIleValPheGluSerCysThrAsnValValAlaSerAsnLeuMetIleAsn 240
DB 711 CAATTCATATCAAAATTTGAGTCACTAATGTTGAGCTTCAAAATTTGATGATCAAT 770
QY 241 AlaSerAlaIleSerProAsnThrAspGluValIleValIleValSerAsnThrGlnTrpIleGln 260
DB 771 GCTTACGCAAAAGGCCCAAAATATCTGATGAGTCCATGATCAAAATCTCAATATTTCA 830
QY 261 IleSerAspThrIleIleGluThrGluAspAspCysIleSerIleValIleSerGluGln 280
DB 831 ATATCTGATCTAATATTGGAACAGATGATGATTTGATTTCAATGTTTCTGATCTCAA 890
QY 281 AsnValGlnAlaThrAsnIleThrCysGluProGluValIleSerIleGluSerIleu 300
DB 891 AATGTGAGGCCCAAAATTTATCTTGCTGTCAGGTCAGTATGATTAAGTATTTGAAAGCTTA 950
QY 301 GlySerGluAsnSerGluAlaIleValIleSerAsnValIleValAsnGluAlaIleIle 320
DB 951 GGATCTGCAAAATTCGAAGCTTATGCTGTAATGTTACTGTAATGAAGCCAAATTAATC 1010
QY 321 GluAlaGluAsnGluValArgIleIleThrTrpGlnIleGluSerGluGlnAlaSerAsn 340
DB 1011 GGTGCCGAAATGAGAGTTAGATCAAGACTTGCGCAGGAGGATCTGCAACAAGCTAGCAAC 1070
QY 341 IleIlePheLeuAsnValGluMetGlnAspValIleValTrpProIleIleIleAspGlnAsn 360
DB 1071 ATCAAAATTTCTGAATGTGAAATGCAAGACGTTAAGTATCCATATTAATTAACCAAAAC 1130
QY 361 TyrCyAspArgValGluProCysIleGlnIlePheSerAlaValGlnValIleAsnVal 380
DB 1131 TATTTGATTCAGTTGAACCATGATTAACAAGTTTTCAGCAAGTTCAAGTGAATAATG 1190
QY 381 ValIleGluAsnIleIleGluThrSerAlaThrIleValIleAlaIleIlePheAspCysSer 400
DB 1191 GTGATGAGAAATATCAAGGCACAAGTCAACAAGGTGCGCATTAATTTGATTTGACAC 1250
QY 401 ThrAsnPheProCysGluGluIleIleIleMetGluAsnIleAsnLeuValGluIleSerGlu 420
DB 1251 ACAAACTTTCCATGGAAGAAATTAATTAATGGAATTAATTAATTAATTAATTAATTAAT 1310
QY 421 LysProSerGluAlaIleThrCysIleAsnValIlePheAsnAsnAlaGluHleValIleThrPro 440
DB 1311 AAACCATCAGAGGCTACGTCGCAAAATGCTATTTTAACAAGTCGACATTTTACACA 1370
QY 441 HisCysThrSerLeuGluIleSerGluAspGluAlaLeuLeuTrpAsnTrp 457
DB 1371 CACTGCACTTCACTAGAAATTTCAAGAGATGAAGCTCTTTGTATTAATTAAT 1421

```

RESULT 3
AX062336      AX062336      1621 bp      DNA      linear      PAT 24-JAN-2001
LOCUS         AX062336
DEFINITION    Sequence 1 from Patent WO0078982.
ACCESSION     AX062336
VERSION       AX062336.1  GI:12540229
KEYWORDS
SOURCE        synthetic construct
ORGANISM      other sequences; artificial sequences.
REFERENCE     1
AUTHORS       Christensen,T.M. and Krelberg,J.D.
TITLE         Process for the enzymatic modification of pectin
JOURNAL       Patent: WO 0078982-A 1 28-DEC-2000;
              DANISCO A/S (DK)
FEATURES
  source       1..1621
               /organism="synthetic construct"
               /mol_type="unassigned DNA"
               /db_xref="taxon:32630"
               /note="cDNA clone"
ORIGIN
Alignment Scores:
Pred. No.:    6,26e-192      Length:    1621
Score:        2390.00       Matches:    457
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match:   100.00%       Indels:    0
DB:            Gaps:        0
US-10-691-374-2 (1-457) x AX062336 (1-1621)
QY      1 MetValIIIGLhAxBhXnsenrIIeLeuLeuIIeIIeIIePhelAAserSerIIeser 20
DB      51 ATGGTTATCCAAAGGAATGATATTCCTTCATTAATTAATTTGGCTTCATCAATTTCA 110
QY      21 ThrCyAsrSerAsnValIIeAspAspAsnLeuPhelVgInValITyrAspAsnIIeLeu 40
DB      111 ACTTGAGAGACAAATGTATTGATGACATTTATTCACAAAGTTATGATATTAATTCCT 170
QY      41 GInGInGluPhelAhiAspPheGInaIaITyLeuSerITyLeuSerITyAsnIIeGlu 60
DB      171 GAACAAAGAAATTTGCATATATTTCAAGCTTATCTTTCTTATTTGAGCAAAATATTTAA 230
QY      61 SerAsnAsnAsnIIeAspLyValAspLyAsnGIIyLeuValIIeAsnValIeuser 80
DB      231 AGCAACAAATATATTCACAAAGTTGATTAATTAATGGATTTAAAGTATTAATGACTTAC 290
QY      81 PheGIIyAlaLyGIIyAspGIIyLysThrITyrAspAsnIIeAlaPheGInGInaIaITyrAsn 100
DB      291 TTTGAGGCTTAAGGGTGATGAAACAAATATGATTAATGACTTTGAGCAAGCATGAAAT 350
QY      101 GIuAlaCySerSerArGrhProValGInPheValIaProLyAsnLyAsnLyAsnITyLeu 120
DB      351 GAAGCATGTTCACTTACAAACACCTGTTCATTAATTTGGTTCCTTAAACAAAGATTAATCTT 410
QY      121 LeuLyGInIIeThrPheSerGIIyProCyAsrSerSerIIeSerValIyIIePheGIIy 140
DB      411 CTCAGCAAAATCACCTTTTCAGGTCCAGTCAGATCTTTATTTTCAGTAAAGATTTTTCGA 470
QY      141 SerLeuGIIuAlaSerSerITyIIeSerAspITyITyAspAspArGaITyITyIIeAlaPhe 160
DB      471 TCCTTAAGAGCATTCATGATAATTTTCAGACTTACAAAGATAGAGGCTTTGGATGACTTTT 530
QY      161 AspSerValGInAsnLeuValIaGIIyGIIyGIIyITrIIeAsnGIIyAsnGIIyValI 180
DB      531 GATAGTGTTCAAATTAATTTAGTGTTCGAGAGAGAGAACTTATCAATGGCAATGGACAAGTA 590
QY      181 TrpITrProSerSerCyLyIIeAsnLySerLeuProCyAsrAspAlaProThrAla 200
DB      591 TGGTGGCCAAAGTTCTTGCAAAATTAATTAATCACTGCATGCAAGGATGCACCAAGGCC 650

```

```

QY      201 LeuThrPheITrPAsnCyLyAsnLeuLyValIaAsnAsnLeuLySerITyAsnIIaGIn 220
DB      651 TTAACTCTTCGAAATTCGAAAAATTTGAAAGGATTAATCTTAAGAGTAAATATGCACAA 710
QY      221 GInIIeAsnIIeLyPheGIIySerCySThrAsnValIaIaSerAsnLeuMetIIeAsn 240
DB      711 CAATTCATATCAATTAATTTGAGTCACCTATATGTTGAGCTTCAAAATTTGATGATCAAT 770
QY      241 AlaSerAlaLySerProAsnThrAspGIIyValIIeValIeSerAsnITrGIIyITrIIeGIn 260
DB      771 GCTTCAGCAAAAGGCCCAATTAATCTGATGAGTCACCTATGATCAATTAATTAATTTCA 830
QY      261 IIeSerAspITrIIeIIeGIIyIThrGIIyAspAspCyIIeSerIIeValIeSerGIIySerGIn 280
DB      831 AATATCTGATACTATTAATTTGAAACAGTGATGATGATTTCAATTTGCTTCGATCTCA 890
QY      281 AsnValGInaIaIThrAsnIIeThrCyGegIIyProGIIyIIeGIIyIIeSerIIeGIIySerLeu 300
DB      891 AATGTCAGGCCCAAAATTAATTAATCTGTGTCCAGGTCAATGATTAATTAATTTGAAAGCTTA 950
QY      301 GIIySerGIIyAsnSerGIIuAlaITyValSerAsnValIThrValIaAsnGIIuAlaLyIIe 320
DB      951 GGATCTGAAATTCGAAAGCTTATATGTGTCTAATGTTACTGTAATTAAGCCAAATTAATTC 1010
QY      321 GIIyAlaGIIuAsnGIIyValaGIIyIIeLySThrITrPInGIIyGIIySerGIIyGInaIeSerAsn 340
DB      1011 GGATCCGAAATTTGAGATGATCAAGCTTGCGAGGAGGATCTGACAAAGCTACGCAAC 1070
QY      341 IIeLySPheLeuAsnValIGIuMetGInAspValIyITyITrPoiIIeIIeIIeAspGInAsn 360
DB      1071 ATCAAAATTTCTGAATGTGGAATTCAGAAATTCAGACCTTAAGTATTCATTAATTAATGACCAAAAC 1130
QY      361 TyrCyAspArGaValGIuProCyIIeGInGInPheSerAlaValaGInaIyLyAsnValI 380
DB      1131 TATGTGATCGAGTTGAACCATGTATCAACAGTTTTCACAGTTCAAGTGAAGAAAATGTG 1190
QY      381 ValITyGInuAsnIIeLySGIIyIThrSerAlaIThrITyValaIaIIeLyPheAspCySer 400
DB      1191 GTGTATGAAATATCAAGGACAGTCACAAAGGTGCGCAATTAATTTGATTTGATTCAGC 1250
QY      401 ThrAsnPheProCySGIIyGIIyIIeIleMetGInuAsnIIeAsnLeuValaGIIyGIIySerGIIy 420
DB      1251 ACAAACTTTTCATGTGAAGGAATTAATTAATGAGAAATTAATTAATTAATTAATTAATTAAT 1310
QY      421 LySPheSerGIIuAlaITrCyLyAsnValIIePheAsnValaGIIyITrPro 440
DB      1311 AAACCATCAAGAGCTACGTCAAAATGTCATTTTAAACATGCTGAACATGTTACACCA 1370
QY      441 HisCySThrSerLeuGIIuIleSerGIIuAspGIIuAlaLeuLeuITyITrAsnITy 457
DB      1371 CACTGCACCTTCATTAAGAAATTTTCAGAGATGAAGGCTTTTGTATTAATTAAT 1421

```

RESULT 4
 LEPG2AR
 LOCUS 1621 bp mRNA linear PLN 09-SEP-2004
 DEFINITION Tomato mRNA for polygalacturonase-2a (PG-2a).
 ACCESSION X04583
 VERSION X04583.1 GI:19291
 KEYWORDS glycoprotein; polygalacturonase; polygalacturonase-2a; signal peptide.
 SOURCE Lycopersicon esculentum (tomato)
 ORGANISM Lycopersicon esculentum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.
 1 (bases 1 to 1621)
 Grierson,D., Tucker,G.A., Keen,J., Ray,J., Bird,C.R. and Schuch,W.
 Sequencing and identification of a cDNA clone for tomato polygalacturonase
 Nucleic Acids Res. 14 (21), 8595-8603 (1986)
 JOURNAL 87066731
 MEDLINE 3786135
 PUBMED

COMMENT
 The predicted N-terminal AA sequence of PG-2a occurs at AA72. Since

the ATG at bp 51 codes for the only inframe methionine preceding the N-terminal sequence of PG-2a, it is probable that the mRNA is translated with a 71aa pre-sequence that is subsequently cleaved to give a 41,828D mature protein.

Data kindly reviewed (22-APR-1987) by D. Grierson.

Location/Qualifiers

1..1621

FEATURES

source

/organism="Lycopersicon esculentum"

/mol_type="mRNA"

/strain="Alisa Craig"

/cultivar="Alisa Craig"

/db_xref="taxon:4081"

/clone="PTOM 6"

/tissue_type="fruit"

/dev_stage="fruit ripe"

51..1424

/note="unnamed protein product; precursor polypeptide (AA

-71 to 386) "

/codon_start=1

/protein_id="CAA28254.1"

/db_xref="GI:19292"

/db_xref="GOA:P05117"

/translation="WVIOKNSILILIIIFASISITCRSNVIDNLFKQVNDILEOF

ADPQAVLSYLSKNTESNNIDKVDKGIKVINLSPGAGGKTYDNIARQAWNEA

CSSTRPVQPVYPRKNKNTLKQITPESCRSSISVYKIPESLSKISDYPKRRMTAF

DSQNLVVGGGITINGQVMPSSCKINKSLPCRDATLTFMCKLKNVNLKSN

AQQLHIKESCTGNVVASNLINAKSPNTDGHVSNQYQISDTILGTDICISV

SGSQNVQATNITCGPHGISIGLSGNSSEAVSNVTNEAKIIGAENGVRKTMQGG

SGQASNIKELNEMQDKYPIIIDONYCDRVPICQPSAVQVKNVYENIKGTATK

VAIKPDGSTNPFCEGIIMENINLVGESGKPSATKQNVFNNAHVTHCHTSLEISD

EALLYNY"

264..1421

/product="mature polygalacturonase (AA 1-386) "

51..263

/note="put. pre-peptide (AA -71 to -1) "

615..623

/note="pot. N-glycosylation site"

768..776

/note="pot. N-glycosylation site"

906..914

/note="pot. N-glycosylation site"

981..989

/note="pot. N-glycosylation site"

1598..1606

/note="pot. polyA signal"

1621

/note="polyA site"

ORIGIN

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-10-691-374-2 (1-457) x LEPG2AR (1-1621)

6.26e-192

2390.00

100.00%

100.00%

8

Length: 1621

Matches: 457

Conservative: 0

Mismatches: 0

Indels: 0

Gaps: 0

1 MetValIIleGlnArgSerIleLeuMetLeuIIleIlePheAlaSerSerIleSer 20

51 ATGGTTATTCAGAAAGAAATGATTCCTCTTCATTAATTAATTTGGCTTCATCAATTTCA 110

21 ThrCysArgSerSerValIIleAspAspAsnLeuPheValIleValIleAspAsnIIleLeu 40

111 ACTTGTAGAGCAATGTTATTCATGACATTTATTCAAACAGATTATATATATTTCTT 170

41 GluGlnGluPheAlaIIleAspPheGlnAlaIleTyrLeuSerTyrLeuSerIleAsnIIleGlu 60

171 GAACAGAAATTTGCTTCATGATTTTCACGCTTATCTTTCTTATTTGAGCAAAATATTTGAA 230

61 SerAsnAsnAsnIIleAspIleValIleAspIleAsnGlyIleIleValIIleAsnValIleuSer 80

221 AGCAACATTAATATGACAGAGTTGATTAATAATGGATTTAAAGTATTAATGACTTAC 290

81 PheGlyAlaIleGlyAspGlyIleValIleAspAsnIIleAlaPheGlnGlnAlaIleTrpAsn 100

291 TTTGGAGCTTAAGGGATGAGTGAATAAATCATATGATTAATATGCTTTGAGCAACATGAAAT 350

101 GluAlaCysSerSerArgThrProValGlnPheValIleProValAsnIleAsnIleLeu 120

351 GAAGCATTTTCATTCAGAACACCTGTTCAATTTGGTTCTTAAACCAAGAAATTCATCTT 410

121 LeuIleValIleThrPheSerGlyProCysArgSerSerIleSerValIlePheGly 140

411 CTCAGCAATTCACCTTTTCAGGTTCACGACATCTTCTATTTTCAGTAAAGATTTTGGCA 470

141 SerLeuGluAlaSerSerIleValIleSerAspTyrIleAspArgArgLeuTrpIleAlaPhe 160

471 TCCTTAGAAGCATTCATGAAATTTTCAGACTTCACAAAGATGAGAGCTTTGGATTCCTTTT 530

161 AspSerValIleGlnAsnLeuValIleGlyIleGlyIleThrIleAsnGlyAsnGlyIleVal 180

531 GATAGTGTTCAAAAATTTAGTTGTTGAGAGAGAGGACCTATCTCAATGCAATGCAAGTAA 590

181 TrpTrpProSerSerCysIleValIleAsnIleSerLeuProCysArgAspAlaProThrAla 200

591 TGTGTGCCAAGTCTTTCGCAAAATTAATTAATTAATCTGCCATGACGAGGATGCACCAACGCC 650

201 LeuThrPheTrpAsnCysIleAsnLeuIleValIleAsnAsnLeuIleSerIleAsnAlaGln 220

651 TTAACCTTCTGGAATTTGCAAAATTTGAAAGGAATTAATCTTAAGAAATGTAATAATGCACA 710

221 GlnIleHISIIleValPheGluSerCysThrAsnValIleAlaSerAsnIleuSerIleAsn 240

711 CAATTCATATCAAAATTTGAGTCATGCACTAAATGTTGAGCTTCAAAATTTGATGATCAAT 770

241 AlaSerAlaIleSerProAsnThrAspGlyValIleIleValSerAsnThrGlnTyrIleGln 260

771 GCTTAGCAAGAAAGCCCAATATCTGATGAGGCCATGATCAATCAATATCAATATTTGCA 830

261 IleSerAspThrIleIleGlyThrGlyAspAspCysIleSerIleValSerGlySerGln 280

831 ATATCTGATTAATTAATTTGAAACAGATGATGATTTGATTTCAATTTGTTCTGATCTCA 890

281 AsnValGlnAlaThrAsnIleThrCysGlyProGlyIleGlyIleSerIleGlySerLeu 300

891 AATGTGACGGCCCAAAATTAATCTTGCTGATGATGATGATGATGATGATGATGATGATGAT 950

301 GlySerGlyAsnSerGluAlaIleTyrValSerAsnValIleValIleAsnGluAlaIleVal 320

951 GGAATTCGAAATTCGAAAGCTTAATGCTTAATGCTTAATGCTTAATGCAATGCAATTAATTC 1010

321 GlyAlaGluAsnGlyValArgIleValIleThrTrpGlnGlyIleSerGlyGlnIleSerAsn 340

1011 GGTCGCAAAATGAGGATTAAGATCAAGACTTGCGGAGGAGATCTGCAACAACTAGCAAC 1070

341 IleIleValPheLeuAsnValIleGluMetCysAspValIleValIleProIleIleIleAspGlnAsn 360

1071 ATCAAAATTTCTGAATGTGAATGCAAGATGCAAGCTTAATGATTCCTTAATTAATCAACAAAC 1130

361 TyrCysAspArgValIleProCysIleGlnGlnPheSerAlaValIleValIleAsnVal 380

1131 TATTTGATTCAGATTGAAACCATGTTATCAACAGTTTTCAGCGATTCAGATGAAAAATGAG 1190

381 ValIleValIleAsnIleIleGlyIleThrSerAlaThrIleValIleAlaIleIleAspPheAspCysSer 400

1191 GGTATGAGAAATATCAAGGGCAACAGTCAACAAAGGTGCGCATTAATTTGATTTGCAAC 1250

401 ThrAsnPheProCysGluGlyIleIleIleMetGluAsnIleAsnLeuValIleGluSerGly 420

1251 ACAACTTTCCATGGAAGAAATTAATGGAATTAATTAATTTAGTGGGGAAGATGGA 1310

421 LysProSerGluAlaIleThrCysIleAsnValIlePheAsnAsnAlaGluIleValIleThrPro 440

Db 1311 AACCATCAGAGCTACGTGCAAAAATGTCATTTTAACAAATGCTGAACATGTTACCA 1370

Qy 441 HieCythrSerLeuGluIleSerGluAspGluAlaLeuLeuTyrAsnTyr 457

Db 1371 CACTGACCTTCACTAGAAATTTGAGAGATGAAAGCTCTTTGTATATATAT 1421

RESULT 5

LOCUS 101809 101809 1624 bp ss-DNA linear PAT 21-MAY-1993

DEFINITION Sequence 1 from Patent US 4801540.

ACCESSION 101809

VERSION 101809.1 GI:269802

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1624)

AUTHORS Hatt, W.R., Sheehy, R.E., Shewmaker, C.K., Kridl, J.C. and Knauf, V.

TITLE Pg gene and its use in plants

JOURNAL Patent: US 4801540-A 1 31-JAN-1989;

Calgene, Inc.; Davis, CA

FEATURES

source location/Qualifiers

1..1624

/organism="unknown"

/mol_type="unassigned DNA"

ORIGIN

Alignment Scores:

Pred. No.: 6.27e-192 Length: 1624

Score: 2350.00 Matches: 457

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

Gaps: 0

US-10-691-374-2 (1-457) x 101809 (1-1624)

Qy 1 MetValIleGlnArgAsnSerIleLeuLeuIleIleIlePheAlaSerSerIleSer 20

Db 47 ATGGTATATCAAGAGATAGTATCTCTTCATTAATTTTGGCTTCATCAATTTCA 106

Qy 21 ThrCysArgSerAsnValIleAspAspAsnLeuPheLeuGlnValTyrAspAsnIleLeu 40

Db 107 ACTGTGAGAGCAAGATGTATATGATGACAAATTAATCAACAGTTATGATATATCTT 166

Qy 41 GluGlnGluPheAlaHisAspPheGlnAlaTyrLeuSerTyrLeuSerIleAsnIleGlu 60

Db 167 GAACAAGATTTGCTCATATTTTCAAGCTTATCTTTTATTTGACAAAATATGAA 226

Qy 61 SerAsnAsnAsnIleAspIleValIleAspIleValIleValIleValIleValIleVal 80

Db 227 AGCAACAAATATATGACAAAGGTGATATAAAATGGAATTAAGATTAATGTAATGCTTGC 286

Qy 81 PheGluAlaIleValIleAspIleValIleValIleValIleValIleValIleValIleVal 100

Db 287 TTTGAGAGCTAAGAGGATGAGAAAACATATGATATATATGATTTGACAAAGCATGGAAT 346

Qy 101 GlnAlaCysSerSerArgThrProValGlnPheValIleProValAsnIleAsnTyrLeu 120

Db 347 GAAGCATGTTATCTAGAAACCTGTTCAATTTGGTCTTCTTAAAAACAAGATTAATCTT 406

Qy 121 LeuIleGlnIleThrPheSerGlyProCysArgSerSerIleSerValIlePheGly 140

Db 407 CTCAAGCAAAATCACTTTCAAGTTCATGCAAGATCTTCTATTCAGTAAAGATTTTGA 466

Qy 141 SerLeuGlnAlaSerSerIleSerAspIleValIleValIleValIleValIleValIleVal 160

Db 467 TCTCTTGAAGAGCTAGTAAATTTTCACTAGCAAAAGATGAAAGCTTTGATGCTTTT 526

Qy 161 AspSerValGlnAsnLeuValIleGlyIleGlyIleThrIleAsnGlyAsnGlyGlnVal 180

Db 527 GATAGTGTCAAATTTAGTTGTTGGAGGAGAGAACTATCATGCAATGACAAAGTA 586

Qy 181 TrpTrpProSerSerCysIleValIleAsnIleSerLeuProCysArgAspAlaProThrAla 200

Db 587 TGGTGCCAAAGTTCTTGCAAATTAATTAATCACTGCCATGACGAGATGCACCAAGGCC 646

Qy 201 LeuThrPheThrAsnCysIleAsnIleValIleAsnIleValIleValIleValIleValIle 220

Db 647 TTAACCTTCTGAAATTTGCAAAATTTGAAAGGAAATTAATCTAAAGATTAATTAATGACAA 706

Qy 221 GlnIleHisIleIlePheGluSerCysThrAsnValIleAlaSerAsnIleMetIleAsn 240

Db 707 CAATTCATATCAAAATTTGAGTCACTATGATGATGATGATGATGATGATGATGATGATGAT 766

Qy 241 AlaSerAlaIleSerProAsnThrAspGlyValIleValIleSerAsnThrGlnTyrIleGln 260

Db 767 GCTTACAGAAAGAGCCAAATTAATCTGATGAGTCCATGATCAATTAATCAATTAATTAATCA 826

Qy 261 IleSerAspThrIleIleGlyThrGlyAspAspCysIleSerIleValIleSerGlySerGln 280

Db 827 ATATCTGATCTATTAATTTGGAACAGTCAATGATGATGATGATGATGATGATGATGATGAT 886

Qy 281 AsnValGlnAlaThrAsnIleThrCysGlyProGlyIleSerIleGlySerLeu 300

Db 887 AATGTCAGAGCCCAAAATTAATTAATCTTGTGTCAGGTCAATGATTAATGATTAAGAGCTTA 946

Qy 301 GlySerGlyAsnSerGluAlaTyrValIleSerAsnValIleThrValIleGlnAlaIleIle 320

Db 947 GATCTGGAATTAATCAAGAGCTTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1006

Qy 321 GlyAlaGluAsnGlyValArgIleIleThrTyrGlnGlyGlySerGlyGlnAlaSerAsn 340

Db 1007 GGTGCGAAATTAATGAGATTAATGATTAATCAACAGTTTGCAGAGGATTTGACAAAGCTTA 1066

Qy 341 IleIlePheIleAsnValIleGluMetGlnAspValIleTyrProIleIleIleIleAspGlnAsn 360

Db 1067 ATCAAAATTTCTGAATGAGAAATGCAAGACGTTAAGATCCATTAATTAATTAATTAATTAAT 1126

Qy 361 TyrCysAspArgValIleGluProCysIleGlnGlnPheSerAlaValIleGlnValIleAsnVal 380

Db 1127 TATTTGATGAGATTAATGAAACATGATTAACACAGTTTGCAGAGCTTCAAGGAAAAAGTG 1186

Qy 381 ValIleGlnAsnIleIleGlyIleThrSerAlaThrIleValAlaIleIlePheAspCysSer 400

Db 1187 GTGATGAGATTAATCAAGGAGCAAGTGCACAAAGGTGCAATTAATTAATTAATTAATTAAT 1246

Qy 401 ThrAsnPheProCysGluGlyIleIleMetGlnAsnIleAsnLeuValIleGlyIleSerGly 420

Db 1247 ACAAACTTTCCATGTAAGAAATTAATTAATGAGAAATTAATTAATTAATTAATTAATTAAT 1306

Qy 421 LysProSerGluAlaThrCysIleAsnValIlePheAsnAsnAlaGluIleValIleThrPro 440

Db 1307 AACCATCAGAGCTACGTGCAAAATGTCATTTTAAACATCTGAAACATGTTTACCA 1366

Qy 441 HieCythrSerLeuGluIleSerGluAspGluAlaLeuLeuTyrAsnTyr 457

Db 1367 CACTGACCTTCACTAGAAATTTTCAAGAGATGAAAGCTCTTTGTATATATAT 1417

RESULT 6

LOCUS A24194 1636 bp DNA linear PAT 04-APR-1995

DEFINITION L.esculentum polygalacturonase clone pTOM6.

ACCESSION A24194

VERSION A24194.1 GI:904404

KEYWORDS

SOURCE Lycopersicon esculentum (tomato)

ORGANISM Lycopersicon esculentum

ECUKaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.

REFERENCE 1 (bases 1 to 1636)

AUTHORS Bridges, I.G., Grierson, D. and Schuch, W.W.

TITLE Anti-sense regulation of plant gene expression

JOURNAL Patent: EP 0532060-A 1 17-MAR-1993;

IMPERIAL CHEMICAL INDUSTRIES PLC; ZENECA LIMITED

FEATURES
source

Location/Qualifiers
1..1636
/organism="Lycopersicon esculentum"
/mol_type="unassigned DNA"
/db_xref="taxon:4081"
51..1424
/codon_start=1
/product="polygalacturonase"
/protein_id="CAA01720.1"
/db_xref="GI:904405"
/db_xref="GOA:P05117"
/db_xref="UniProt/Swiss-Prot:P05117"
/translation="MVIONSTILLIIIPASISTCRSNVIDNLPKOVYNIIEOF
AHDPAVYLSKINIESNNIDKVDKNGIKYINYSFPAKGGKTYNDINAEQANNEA
CSSRTVPVPEKKNYILKQITFSGPCSSISVKI PSLSESSKISDYKDRRLTAF
DSVNLVVGGGTNGNGQVWPPSSCKINSLPCRDAPALTPMCKLKNVNLKSN
AQOIHKEPSCITNVVASNLIMINAKSPEDDVSNVOYIOISPTIIGTGDCISIV
SGSONVANTICGPHGISIGLSGNSSEAVSNVTNEAKIIGAENGVRKTMGG
SGASINKEINVMQDVKPIIIDQNYCDRVPKCIQPSAVQKRVNVENTKIGSATK
VAIKDCSTNPFCEGIIMENINLVESGKPEATCKNVHFNNAEHVTHCTSLISED
EALYNY"

ORIGIN

Alignment Scores:
Pred. No.: 6,336-192 Length: 1636
Score: 2390.00 Matches: 457
Percent Simlarity: 100.00% Conservative: 0
Best Local Simlarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-691-374-2 (1-457) x A24194 (1-1636)

Qy 1 MetValIIEGInARgAnSerIleuLeuLeuIleIleIlePheAlaSerSerIleSer 20
Db 51 ATGGTTATCCAAAGGAATGATATTCCTTCATATTAATTTTGCTTCATCAATTTCA 110
Qy 21 ThrCyARgSerARnValIleAhpAhpAenIleuPheIleuValIleAhpAenIleu 40
Db 111 ACTTGTAGAACCAATGATTAATGATGACAAATTTATCAACAGTTATATATATTTCTT 170
Qy 41 GIuNGInIuPheAlAHIsAPheGInAlATyrLeuSerTyrLeuSerTyrAenIleu 60
Db 171 GACCAAGAAATTTGCTCATGATTTTCAAGCTTATCTTTCTTATTGACAAAATTTGAA 230
Qy 61 SerAenAenAenIleAhpIleuValIleAhpIleuValIleAhpValIleAhpValIleuSer 80
Db 231 AGCAACAATAATATGACAAAGGTTGATAAATGGGATTAAGTATTATGACTTAC 290
Qy 81 PheGIyAlAlYyGlyARgIlyIySerThrTyrAhpAenIleAlaPheGIuGInAlATrPheN 100
Db 291 TTTGAGCTTAAGGGGATGGAATAAACTATGTATTAATTCATTGACCAACATGGAAAT 350
Qy 101 GIuAlaCySerSerArGthrProValGIuPheValIleProIyAenIyAenIyLeu 120
Db 351 GAAGCATGTTCTATGAAACACCTGTTCAATTTGGTTCCTTAAACAAAGAAATTTCTT 410
Qy 121 LeuIyGInIleThrPheSerGIyProCyARgSerSerIleSerValIyAenIlePheGIy 140
Db 411 CTCNAGCAATCACTTTTCAGCTCCATGCATGCTTCTTATTTCAGTAAAGATTTTGGAA 470
Qy 141 SerLeuGIuAlaSerSerIyAlleSerAhpIlyIyAhpAhpArGLeuTrIleAlaPhe 160
Db 471 TCCTTAGAAGCATCTAGTAAATTTTCAGCTCAAGAAATGAAGCTTGGATTCCTTTT 530
Qy 161 AhpSerValAlaAhpLeuValIleGIyGIyGIyIleThrIleAhpGIyAenGIyGInVal 180
Db 531 CATAGTGTTCAAATTTTACTTGTGGAGAGAGAACTATCAATGGCAATGGACAAAGTA 590
Qy 181 TrpTrpProSerSerCyAlYyAlleAenIySerLeuProCyARgAhpAlaProThrAla 200
Db 591 TGGTGCCCAAGTTCTTGCAAATTAATAATCATGCGCAATGCAAGGATGCCAACAGGCC 650

Qy 201 LeuThrPheTrpAenCyAlYyAenIleuIyValIleAenAenIleuIySerIyAenAlaGIn 220
Db 651 TTAACCTTCTGGAATTCGAAATTTGAAATTTGAAATTTCTTAAAGTAAGTAATGACACAA 710
Qy 221 GIuIleAlaIleYyPheGIuSerCyArThrAenValIleAhpAenIleuAenIleAen 240
Db 711 CAATTCATATCAATTAATTTGACATGACATATGTTGATGCTTCAATTTGATGATCATAT 770
Qy 241 AlAserAlaYySerProAenThrAhpGIyValIleAhpValIleAhpThrGIuTrIleGIn 260
Db 771 GCTTAGCAAGAACGCCAAATATCTATGATGAGTCCATGTATCAATTAATTTCTCA 830
Qy 261 IleSerAhpTrIleIleGIyThrGIyAhpAhpCyAlleSerIleValSerGIySerGIu 280
Db 831 ATATCTGATACATATTTAGAAACAGTATGATGATTTCAATTTCTTCTGATCTCA 890
Qy 281 AenValGIuAlaThrAenIleThrCyGIyProGIyAlaGIyIleSerIleGIySerLeu 300
Db 891 AATGCGAGGCCACAAATTTTACTTGTGTCAGGTCAAGTCAATGGTATTAAGTATGGAGCTTA 950
Qy 301 GIySerGIyAenSerGIuAlaTrValSerAenValIleThrValAhpGIuAlaYyIle 320
Db 951 GGATCTGGAATTCGAAAGCTTATGCTATGCTATATGTTACTGTAAATGAAGCCAAATTTATC 1010
Qy 321 GIyAlaGIuAenGIyValArGIyIleYyThrTrpGIuGIySerGIyGInAlaSerAen 340
Db 1011 GGTGCCGAAATGAGTTAGTATGATCAAGCTTGCGAGGAGATCTGACACATGACAC 1070
Qy 341 IleYyPheLeuAenValGIuAenIleAhpValIyYyTrpProIleIleIleAhpGIuAen 360
Db 1071 ATCAATTTCTGAATGATGGAATGCAAGCGTTAAGTATCCATATTTATTAACCAAAAC 1130
Qy 361 TyrCyAhpArGValGIuProCyAlleGIuIlePheSerAlaValGIuValIyAenVal 380
Db 1131 TATGTGATCGAGTTGAAACATGTATACAAAGTTTCAGCGTTCAAGTCAAAATGTG 1190
Qy 381 ValTyrGIuAenIleYyGIyThrSerAlaThrIyAenValIleIleYyPheAhpCySer 400
Db 1191 GTGATGAAATATCAAGGCGACAGTGCACAAAGGTGGCCATTAATTTGATGTCACAC 1250
Qy 401 ThrAenPheProCyGIuGIyIleIleAenGIuAenIleAenIleuValGIyIleSerGIy 420
Db 1251 ACAACCTTCCATGGAAGGAATTTAATGAGAAATTAATTTAGAGGAAAGTGA 1310
Qy 421 LybProSerGIuAlaThrCyAlYyAenValIlePheAenAenIleGIuIleValIleThrPro 440
Db 1311 AAACCATCAAGAGGCTACGCAAAATGTCATTTTAACAATGCTGAACATGTTACACCA 1370
Qy 441 HisCyTrpSerLeuGIuIleSerGIuAhpGIuAlaLeuLeuThrAenTyr 457
Db 1371 CACTGCACTTCACTAGAAATTTCAAGAGATGAAAGCTTTGTATATATAT 1421

RESULT 7
AR364905 1636 bp DNA linear PART 03-SEP-2003
LOCUS
DEFINITION Sequence 2 from patent US 5447867.
ACCESSION AR364905
VERSION AR364905.1 GI:34428082
KEYWORDS
SOURCE
ORGANISM
Unknown.
REFERENCE
1 (bases 1 to 1636)
Bridges, I., Schuch, W. and Grierson, D.
TITLE Recombinant DNA containing pectin esterase gene segments
JOURNAL Patent: US 5447867-A 2 05-SEP-1995;
FEATURES
source
1..1636
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Alignment Scores:


```
/translation="MVIORNSILLIIIFASSISICRSNVID)NLFKOVNDILIEOF
AHDROAYLSYKIESNNNI DKUDKGIKXINULSFGAKDGKTYNDIAEEOANMEA
CSSTPVOFVPRKNKNTLLKQITRSGPSSISVKIFESLAA;SKISDYPKORRLMIAP
DSVONLVGGGGLINGQVWVPSSCKINKSLPCDAPLALTFMCKKLKYNLKSXN
AQOQHIKPEKCTNVASNMIMNASAKSPNTGIVHSNTOYIOISDTIIGTDDCISIV
SGSONVOAKNTICGPHGISIGLSGNSBAVSVNVTNEAKIIGAENGVILKTMQGG
SGOASNIKEIENVMODVXKPIIIDONCYDRVPECIQOPSADVONVVENIKGMSATK
VAIKPDCSTNPECGIIMENINLVGSEKSPSEATCKNHPNNAEHVTHCTSLSESD
EALXNY"
intron 1758..2415
      /number=1
exon 2416..2547
      /number=2
intron 2548..3326
      /number=2
exon 3327..3491
      /number=3
intron 3492..3695
      /number=3
exon 3696..3716
      /number=4
intron 3717..4259
      /number=4
exon 4260..4467
      /number=5
intron 4468..4566
      /number=5
exon 4567..4648
      /number=6
intron 4649..5601
      /number=6
exon 5602..5710
      /number=7
intron 5711..6138
      /number=7
exon 6139..6255
      /number=8
intron 6256..6787
      /number=8
exon 6788..7244
      /number=9
polyA_site 7244
ORIGIN
Alignment Scores:
Pred. No.: 1,31e-116 Length: 7456
Score: 1506.00 Matches: 454
Percent Similarity: 24.43% Conservative: 0
Best Local Similarity: 24.43% Mismatches: 3
Query Match: 63.01% Indels: 1404
DB: 8 Gaps: 8
US-10-691-374-2 (1-457) x LEPOLYGA (1-7456)
QY 1 MetValIIeGInARgAsnSerIIeLeuLeuIIeIIeIIePheAlaSerSerIIeSer 20
Db 1479 ATGGTTATCCAAAGGAATAGTATTCCTCTTCATTAATATTGCTTCATCAATTTCA 1538
QY 21 ThrCyASerSerAsnValIIeAsPAsnLeuPheLyGInValIlyrAspAsnIIeLeu 40
Db 1559 ACTTGTAAAGCAAGTATTAATGATGACAAATTTATCAACAGTTTATGATAATTTCTT 1598
QY 41 GIUGInIuPheAlIHLeAsPheGInAlAlYrLeuSerIYrLeuSerIyAsnIIeGlu 60
Db 1599 GAACAAAGATTTGCTCATGATTTTCAAGCTTATCTTTCTTATTTGAGCAAAAATATTGA 1658
QY 61 SerAsnAsnAsnIIeAsPlyAsnValIAsPlyAsnGlyIleValIIeAsnValIIeSer 80
Db 1659 AGCAACAATTAATATTGACAAAGGTGATTAATAAAGGATTAAGATTAATGATTAAC 1718
QY 81 PheGInAlAlYrGlyAsPQIlyrThrIYrAspAsnIIe----- 93
Db 1719 TTTGAGGCTAAGGATGAGAAABACATATGATTAATATTGTAAGTATTTAAATATTGGA 1778
```

```
QY 93 ----- 93
Db 1779 TATATTGTGGGAGTGAATAATGATAGATAATTAAGATTATTGGAAAGATGAAGATT 1838
QY 93 ----- 93
Db 1839 ATATTATTAAGTAAGAAATTAATTTTCTCGTTTGTAGATTAAAGTGAAGATGATTTC 1898
QY 93 ----- 93
Db 1899 TCGTTAAGCAGAGAAAGCTATTTTCCATGGTAAGTATTTTCTTACTTTTAATTA 1958
QY 93 ----- 93
Db 1959 CGTCATGATTTTGTCTATTAAGTCAAGATTAAGACACTATTATTGATGATTAAGTCTCGA 2018
QY 93 ----- 93
Db 2019 AAAGAAATGATAGTAATTTTGTCTTAATATACTAATTTCTTATATGATATTTTTC 2078
QY 93 ----- 93
Db 2079 AACCAAAATTAACAAAGCGTAATCCATTAAGTGGCGCTAGAAATTAAGTAAGTTCTAT 2138
QY 93 ----- 93
Db 2139 TCAATTTTAACCTTATTTAATTTTAGTGAAACCTCGACAAAACGAACAAAGTATTC 2198
QY 93 ----- 93
Db 2199 AAACCTTTATATTCGGAATTCGAGACCAACATATGAACAACCTGACACATGATAGT 2258
QY 93 ----- 93
Db 2259 CCTAATATATTAATTTTCTTAATAAATATCTTCAATCTACCATATGAAATATTGAAA 2318
QY 93 ----- 93
Db 2319 ATGACTTTATTCATTCGACACATATCAAGAGTTTCTTTAAGAATTTACCACTACAT 2378
QY 94 ----- -AlaPheGIUGInAlaIrpAsnG 101
Db 2379 TTGATATGTTCTTATCGGTTAATAATATATCTTCAGGATTTGAGCAAGATGAAATGA 2438
QY 101 uAlaCySerSerArgIhrProValGInPheValIAlProLyAsnLyAsnIYrLeuLe 121
Db 2439 AGCATGTTCACTAGAACACCTGTTCAATTTGTGTTCTTAATAAACAAAGAAATTAATCTTCT 2498
QY 121 uLyGInIleIhrPheSerGlyProCyArgSerSerIIeSerValIyS----- 137
Db 2499 CAAGCAAAATCACCTTTTCAGGTCATGACAGATCTTCTATTTCAGT-AAAGGTTAGCATAT 2557
QY 137 ----- 137
Db 2558 TGATTTATTAATCTCTTTGTGTAGCAATATATTAATCTGTTTATGACAAAATTTAAGA 2617
QY 137 ----- 137
Db 2618 AGTATCAAGATAGATTAACAATGAATTTTGTCACTAATTTAGCGGATTAAGTGAAGAA 2677
QY 137 ----- 137
Db 2678 TTATCAAAATGTATAGTATGATGACAACTTAGCTATGAATTTAGCTAGTGAAGAGTT 2737
QY 137 ----- 137
Db 2738 TGATCTAATTTCTATTTTCTTTTGTAGAGTAAGATATTGAAAACATGTATTAATTA 2797
QY 137 ----- 137
Db 2798 TTAATTAATGCTTAATTAATATATGCAATGATAGTTCAAACTTAAGAACTGTCAAAAGAA 2857
```

```
QY 137 ----- 137
Db 2858 AATAAGAAAGAAATATTTTATTTTAAATAAATAAAGAAAAATGAGAAATTAATT 2917
QY 137 ----- 137
Db 2918 CAAACGAGAGATTAATCAATATCTATGCGGATTAAGATATATATATATGTAAGAAA 2977
QY 137 ----- 137
Db 2978 CAGCACTACACATATCTAATAAGTCTCATTAAGATATATAAATAAGTGTAGCAAA 3037
QY 137 ----- 137
Db 3038 CAGTTATCCCTACAAAACTTTGTGGGTAGATGATCCAGAGTTGTTCCAGACTCT 3097
QY 137 ----- 137
Db 3098 TGCTTAATAAAATGTTTTTCTTAATAAGTTGAAAGAAATGTATATGATGAAAAATAT 3157
QY 137 ----- 137
Db 3158 GAAAGAAAAACATATCAATATTAATAAATTAAGTAATCAAGTAAACGAATAACAAATA 3217
QY 137 ----- 137
Db 3218 GGAATAATCTCATTAATAAGAAATTTAGTGTTCGTTAACAATATCTTAGTTATTC 3277
QY 138 ----- 141
Db 3278 ATTGTTCTTAAATTCCTTCCTTATTTTGAATTAATCTAATGCAAGATTTTGGATC 3337
QY 141 ----- 161
Db 3338 CTTAAGAGCTCTAGTAAATTTTCAACATCAAAAGATAGAGCTTTGATTCCTTTCA 3397
QY 161 ----- 181
Db 3398 TACTGTCAAAATTTGTTGTTGAGAGAGAGAACTATCAATGCAATGCAAGATATG 3457
QY 181 ----- 192
Db 3458 GTGGCCAAAGTTCTTGCAAAATTAATAATCACTGGTAATTTAATACCTTGCTTAAGT 3517
QY 192 ----- 192
Db 3518 TTACGCTATGTGCTCGAATTTCTTAACTTCTAAGATATATATTTGAAGA 3577
QY 192 ----- 192
Db 3578 GGTGTCACAAAATGCATCATTTTAAAGATTCGACCAATATTAAGTTTATGTAATCTA 3637
QY 193 ----- 193
Db 3638 ATTTTCAGAGCATCTTTCCTCTGTAATCATTTGATACCTTTTCTTCATGACGCC 3697
QY 193 ----- 199
Db 3698 ATGCGAGATGCAACCAACGTACGTAAATTCATTTGATTAATAAAAAAAGCCTA 3757
QY 199 ----- 199
Db 3758 AAATATATTTGAATTTAATGAAGGTAAATTAATTTTAACCTTGGCAGAGACCTAT 3817
QY 199 ----- 199
Db 3818 TACCCCTTGACATTTAATAGTATTTTAAAGATTAATAAGTGTAGTGAACAA 3877
QY 199 ----- 199
Db 3878 AATTAGATATCAAAAACTATTTGAATAATTAATTAATGAATTTTTCATATCA 3937
QY 199 ----- 199

QY 137 ----- 137
Db 3938 ATATGATTAATAAAATATAGTTAAAGTCTTATGATTTGATTTCAAAAAATAAATCATG 3997
QY 199 ----- 199
Db 3998 ACMAACAAATAGTAGACGGAAGAAATATTAACATACCTCTTCAAGTAGAATTCATTTGT 4057
QY 199 ----- 199
Db 4058 AACACACCTCAAAAACCTACGTTTCTTGATTAATATTCCTTAAATAGTAA 4117
QY 199 ----- 199
Db 4118 TCAAAAGCTATATAGTCTGTCAAAATCTATACATTGGAACCTATCTTGACGCCCTGT 4177
QY 199 ----- 199
Db 4178 ACATTGAGATCGTTGAACAAATGATGATTAATTAACCTTGTATTTAAATATTTAA 4237
QY 200 ----- 212
Db 4238 ACTAATATTTGTTAATTTTTCAGGCTTACCTTCTGAAATTCGAAATTTGAAGTGA 4297
QY 212 ----- 232
Db 4298 TAACTTAAAGGTAAAAATGACAAACAAATTCATCAATTTGAGTCATGACACTATGT 4357
QY 232 ----- 252
Db 4358 TGTAGCTTCAATTTGATGATGATCAATGCTTACGAAAGGCCCAAAATCTGATGAGATCA 4417
QY 252 ----- 268
Db 4418 TGTATCAAAATCAAAATATATTCAAATATCTGATCATATATGAACAGTTATTAAT 4477
QY 268 ----- 268
Db 4478 TTAATTTTATTTATCCAAATTAATTAAGAAAAAAGAGATTTTATTTGATTAATCA 4537
QY 269 ----- 279
Db 4538 AATTATTAATTTTAAATTTTATTTTATAGTATGATTTGATTTCAATTTCTTGATC 4597
QY 279 ----- 296
Db 4598 TCAAAATGTCAGGCCCAAAATTAATTAATCTGTGTCTCAGGTCAATGATTAAG-CTACTCTA 4656
QY 296 ----- 296
Db 4657 TTTTCAAAATATATCTGTTTCATTTTCTATTTTCAATAAAGTATGATTAATTA 4716
QY 296 ----- 296
Db 4717 TTAATTTAATCTTTAATTAATTAATTTATGCAAAATTTTCTGTGCTTATGTTAT 4776
QY 296 ----- 296
Db 4777 GACTTAGCAATATATAGGCGCTTGGATGGCGAATTAAGACCTTTAAAAAGTA 4836
QY 296 ----- 296
Db 4837 CTTTAAAGGTGTAATCTTATTTTAAATAAAGAGTATCGTTTGAATAAAGTGC 4896
QY 296 ----- 296
Db 4897 TGAAGTTGTATGTCAACGTAAGGAAGGAAGAAATGGAAGAAAGTGTAGGTTATA 4956
QY 296 ----- 296
Db 4957 TGGGTATTTGTATTAATAATTAATTAAGACAAAAAGATTAATAATGTGTCAACTTAACA 5016
QY 296 ----- 296
```

Db 5017 ACTTAAAGCTACCTACCCCTACCCAGCTTTAACTTTGGCTTAAATTAAGTTT 5076
QY 296 ----- 296
Db 5077 TTTTAAAACTTAAATAAGTTGTTTGAGTATTCGCCAAGAGCTAAATATGCAAAAACC 5136
QY 296 ----- 296
Db 5137 AGCTTTAAGTCAGTTTGACCAAGCTTTTAAAGTGAAGCCAAAGAGCTTTAAATAGTCTG 5196
QY 296 ----- 296
Db 5197 CTTAGATGCTATATATATTTGAGCTTTTGAAGTAGTATATTCCTTAAGTTCAA 5256
QY 296 ----- 296
Db 5257 CATAAATAATACATGCTTTAATACATAGCAATATAGTTAATCAAAAGACGAATGATGATTA 5316
QY 296 ----- 296
Db 5317 TTTTGGCAATTTGATTTATTCACAGAAAAGGATAGTTCAAAAGTACATTCATGAAT 5376
QY 296 ----- 296
Db 5377 TGAAGATATCAATAAAGACTAAATAATTAAGAAATCAATATGAGGATCAAAAATGTTAT 5436
QY 296 ----- 296
Db 5437 TACCTTATTAATAATCTATTCATTTTCATATTAATTAATTAAGAGTTTATA 5496
QY 296 ----- 296
Db 5497 ATCTAATAAACAATCAATATTAATTAATGCAAAATGTTTGGTACCTATATCTTCT 5556
QY 297 ----- 301
Db 5557 GAATATTTGCTCTATTTTCTCTTTTATTTTCCATGAGATTACTATGGAAGCTTAGG 5616
QY 301 ySerGIyAenSerGIuAlaTYrValSerAenValThrValAenGIuAlaYelIleGI 321
Db 5617 ATCTGAAATTCAGAAAGCTTATGTCTATGTCTACTGTAAATGAAGCAAAATTAATCGG 5676
QY 321 yAlaGIuAenGIyValAenGIyIleYThrPglIn----- 332
Db 5677 TGCCGAAATGAGTTTGAATCAAGACTTGCA-GGTACCTCCCCCCCCCCCCCCC 5735
QY 332 ----- 332
Db 5736 CACAGGCCCATTTTAAATTTTAAATTTTAAATTTTGAATATCAATATTAAGATTAA 5795
QY 332 ----- 332
Db 5796 ATTGATTTTCATGTTGAATTTTAAATTTTGAATTAAGATATGATATTTTCTAGCTTCTAT 5855
QY 332 ----- 332
Db 5856 GTTATATAGAAAAAAATGTTCAAGAACTTCAGATTATTTGACTGTACTAAGTGAAT 5915
QY 332 ----- 332
Db 5916 GTGTGCTTTGTTTGAAGTTTGGTTTATCCAGTTTGGGTGATGATTAACCAACTTA 5975
QY 332 ----- 332
Db 5976 TAATGAAAAAGGGCTGCAACGGCCGACCTAGTCTAGTATCAATAGAAAGATCTCAC 6035
QY 332 ----- 332
Db 6036 GTCTGTTATTCAGATGACGTTCTTGTTGAATGTTAATTAATTAATTAATTAACA 6095
QY 333 ----- 338
Db 6096 TGTATTAAGCATATATATAATTAATGCTGTTTAAATGCTAGGAGATCTGGACAGC 6155

QY 338 aSerAenIleYSerPheLeuAenValGIuMeGIuAenValYerYrProIleIleLeAs 358
Db 6156 TAGCAACATCAAAATTTCTGAATGTGAATGGAAGACGTTAAGTATCCATATTAATTAAGA 6215
QY 358 pGIuAenTYrCyAaPArGValGIuProCyAeGIu----- 370
Db 6216 CCAAAACTTGTGTGATCGAGTTGAACATGTATACACAGGTATTTTATTAACGAAC 6275
QY 370 ----- 370
Db 6276 AATTATTATATTTTATTACTTCTTAATACACTTACATCATTAATAAATTGAGATCTT 6335
QY 370 ----- 370
Db 6336 TTCCTAGTTAGTAACTTTTGAATAGATTTTGAATATGATATTCATTAATCCTTTTA 6395
QY 370 ----- 370
Db 6396 TTTTCTCTAATTTATGATCTTTTGGACATATGGCTTAATAATCTGTAAAGTAAACT 6455
QY 370 ----- 370
Db 6456 GAATATCATTAAGAAAAATGTTAGATTAATCTAAATTTTATTAATTAATTAAGAGTT 6515
QY 370 ----- 370
Db 6516 ATCTAATATTTTGTATGTAGATTTAGAAACATATATACATTAATAACATTAGATTCAAATTTA 6575
QY 370 ----- 370
Db 6576 ATTAATATCAAAATTTGATTCAATCAATCAATGACTACACAAAGAAATACATGCAGATT 6635
QY 370 ----- 370
Db 6636 CTAACAATATGATGAAGTCAATTTCAAAACGAATCAATATATAGATATATCCTTAA 6695
QY 370 ----- 370
Db 6696 AAGAGCATTTGGTAAATAAGTAAATATCAATTAAGTTATTAATAAAATTTCAACTGCA 6755
QY 371 ----- 381
Db 6756 TCTCTCAGATTATTTAATCACTTGTTCAGATTTTTCAGCAATTCAGGAAAAATGTGG 6815
QY 381 aTYrGIuAenIleYSerGIyThrSerAaThrLyAenValaIleYSerPheAerCySerT 401
Db 6816 TGTATGAGAAATATCAAGGCAACAGTGCACAAAGGTGGCATTAATAATTTGATTGCACGA 6875
QY 401 hAaenPheProCyAeGIuGIyIleIleMeGIuAenIleAenLeuValaGIyGIuSerGIyL 421
Db 6876 CAATCTTTCCATGATGAAGAAATTAATGAATTAATTAATTAATTAATTAATTAATTAAT 6935
QY 421 yEProSerGIuAlaThrCyAeLyAaenValaIlePheAaenValaGIuIleValaIleThrProH 441
Db 6936 AACCATCAGAGCTACGTGCACAAAAATGTCATTTTAACATCTGAACATGTTACACAC 6995
QY 441 IeCyerThSerLeuGIuIleSerGIuAerGIuAerGIuAerGIuAerGIuAerGIuAerGIu 457
Db 6996 ACTGCATTCACATGAATAATTTCAAGAGATGAAGTCTTTGTATATATAT 7045

RESULT 9
TOMPGAA
LOCUS TOMPGAA 7456 bp DNA linear PLN 27-APR-1993
DEFINITION Tomato polygalacturonase (PG) gene, exons 1-9.
ACCESSION M37304
VERSION M37304.1 GI:170472
KEYWORDS polygalacturonase.
SOURCE Lycopersicon esculentum (tomato)
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotids;
asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.

REFERENCE 1 (bases 1 to 7456)
AUTHORS Bird,C.R., Smith,C.J., Ray,J.A., Moureau,P., Bevan,M.W., Bird,A.S.,
HUGHES,S., Morris,P.C., GRIERSON,D. and Schuch,W.
TITLE The tomato polygalacturonase gene and ripening-specific expression
JOURNAL in transgenic plants
COMMENT Plant Mol. Biol. 11, 651-662 (1988)
ORIGINAL source text: Tomato (strain M11 cv Alisa Craig) DNA,
clone pTOM6.
FEATURES
source
1..7456
/organism="Lycopersicon esculentum"
/mol_type="genomic DNA"
/db_xref="taxon:4081"
prim_transcript 1411..>7456
/note="polygalacturonase mRNA and introns"
CDS
join(1479..11757,2416..2547,3327..3491,3696..3716,
4260..4467,4567..4648,5602..5710,6139..6255,6788..7048)
/note="polygalacturonase"
/codon_start=1
/protein_id="AA34178.1"
/db_xref="GI:170473"
/translation="MVYQNSILLLIIIFASSISTCKSNVIDNLEKQVYDNIIEQEF
AHDROAVLSTSKIESNNINIDKVDKNGIKVINVLSPGAGDKETDNLAEQANMEA
CSSRTPVQFVPEKKNKYLKQITFSGCRSISVIKIIGSLASKSIDYKDRLLMTAF
DSVNLVVGGGGTINGGVWBPSSCKINSLPCRDAPLPTFMKNKNNLKKN
AQQIHIFESCTNVASNLIMINASAKSPNTDGVHNTQYIOISDTLIGGDCISIV
SGSQNVATNTICGPHGISIGSLSGNSPAAVSVNTVNAKIIQENGRIRKTVGG
SGQSNIKPLNVEMQDVKPIIIDQNCDEPCTIOQPSAVQKVVYENIKGTSAIK
VAIKFDCSTNPCEGIMENINIVGSEKFSKATCKNVHNAHTPCTSLSESD
EALLVNY"
exon <1479..11757
/note="polygalacturonase"
/number=1
intron 11758..2415
/note="polygalacturonase intron A"
exon 2416..2547
/number=2
intron 2548..3326
/note="polygalacturonase intron B"
exon 3327..3491
/note="polygalacturonase intron C"
intron 3492..3695
/note="polygalacturonase intron C"
exon 3696..3716
/note="polygalacturonase intron D"
intron 3717..4259
/note="polygalacturonase intron D"
exon 4260..4467
/note="polygalacturonase intron E"
intron 4468..4566
/note="polygalacturonase intron E"
exon 4567..4648
/number=6
intron 4649..5601
/note="polygalacturonase intron F (no splice consensus);
putative"
exon 5602..5710
/number=7
intron 5711..6138
/note="polygalacturonase intron G"
exon 6139..6255
/note="polygalacturonase intron H"
intron 6256..6787
/note="polygalacturonase intron H"
exon 6788..>7048
/note="polygalacturonase"
/number=9
ORIGIN
Alignment Scores:
Pred. No.: 1,31e-116 Length: 7456
Score: 1506.00 Matches: 454
Percent Similarity: 24.43% Conservative: 0
Best Local Similarity: 24.43% Mismatches: 3
Query Match: 63.01% Indels: 1404
DB: 8 Gaps: 8

US-10-691-374-2 (1-457) x TOMPGAA (1-7456)
QY 1 MetValIleGlnArgAsnSerIleLeuLeuIleIlePheAlaSerSerIleSer 20
Db 1479 ATGGTATCCAAAGAAAGTATTCCTTCCTTCATTTATTTTGGCTCACTCAATTC 1538
QY 21 ThrCysArgSerAsnValIleAspAspAsnLeuPheValIleValTyrAspAsnIleLeu 40
Db 1539 ACTGTGAGAGCAAGCAATGTATGATGACATTTATTCAAACAGTTATGATATTTCTT 1598
QY 41 GluGlnGluPheAlaHisAspPheGlnIleTyrLeuSerTyrLeuSerIleAsnIleGlu 60
Db 1599 GAAACAAGATTTGCTCATATTTTCAAGCTTATCTTCTTATTTGAGCAAAAAATATTTGAA 1658
QY 61 SerAsnAsnAsnIleAspIleValIleAspIleAsnGlyIleValIleAsnValIleAsn 80
Db 1659 ACCCAACAAATTAATTAAGCAAGGTTGATTAATTAAGGATTAATTAAGTATTAATGCTT 1718
QY 81 PheGlyAlaIleGlyAspGlyIleThrTyrAspAsnIle----- 93
Db 1719 TTGGAGCTTAAGGATGATGCAAAACATATGATATATTTGAAGTATTAATATTTGAAA 1778
QY 93 ----- 93
Db 1779 TATATTTGTGGGATGAATAATGATAGAAATATTAAGATTTTGGAAAGATGAAGATT 1838
QY 93 ----- 93
Db 1839 ATATTTTATAAGTAAGAAATTAATTTTCTGTTTATAGTATTAAGTGAAGAAATGAGTTTC 1898
QY 93 ----- 93
Db 1899 TCGTTAAGCAGAAAGATATTTTCCATGATAGTATTTTCTTATTTTAAATA 1958
QY 93 ----- 93
Db 1959 CGTATAGTATTTGCTATCTCAAGATTAAGACATTAATTAATTAAGATTTAGTCTCGA 2018
QY 93 ----- 93
Db 2019 AAGAAATGATATGTAATTTTGTCTTAATTAATTAATCAATTTCTTAATTAATTTTTC 2078
QY 93 ----- 93
Db 2079 AACCAAAATTAACAAGCTTAATCAATAAGTGGCTCTAGATAAAGATTAAGTTCTAT 2138
QY 93 ----- 93
Db 2139 TCAATCTTAACCTTATTTATTTATTTAGTGAACCTCGACAAAACAAACGATATTC 2198
QY 93 ----- 93
Db 2199 AAACTTTATATTCGAATTCAGACCAACATATGAAACAACCTGACATGATATAGT 2258
QY 93 ----- 93
Db 2259 CCTAATATATATTTTCTTAATAAATAATCTTAATCAATCAATTAATTAATTAATA 2318
QY 93 ----- 93
Db 2319 ATGACTTTATCTATTCGACACATATCAAGAGTTCTTTAAGATTTTACCACTATCAT 2378
QY 94 -----AlaPheGluGlnAlaIleTyrPheGlu 101
Db 2379 TTGATATGTTCTTATCTGTTTAAATTAATTTTCAAGCAATTTGACAGCATGCAATGA 2438
QY 101 UAlaCysSerSerArgThrProValGlnPheValIleProIleAsnIleAsnTyrLeuLeu 121
Db 2439 AGCATGTTCACTGAGAACACCTGTTCAATTTGTTGTTCTTAATAAACAAGATTTCTTCT 2498
QY 121 UlyGlnIleThrPheSerGlyProCysArgSerSerIleSerValIys----- 137
Db 2499 CAAACAATCACCTTTTCAAGTCCATGACATCTTCTATTTTCAGT-AAAGTTAGCATAT 2557

QY 137 ----- 137
Db 2558 TGATATTATTAATCCTTTGTAGCAATATATATCTGGTTATGACAAATTAAGAA 2617
QY 137 ----- 137
Db 2618 AGTAATCAAGATAGATAAACAATGATTTTGTCTACTTAATTTAGCGGATTAGTGAGAA 2677
QY 137 ----- 137
Db 2678 TTATCAAAATGTATGTAGTATAGCAACTAGCTATGAATAGCTAGTAAGAAGTT 2737
QY 137 ----- 137
Db 2738 TGATCTAATCTATTTTTTTTTTGTAGAGTAAGATATTTGAAAACATGTATTAATTA 2797
QY 137 ----- 137
Db 2798 TTAAATTAGTCTTAATTAATATATGTCATGATGATTTCAAACTAAGAACTGTCAAAAGAA 2857
QY 137 ----- 137
Db 2858 AATAAGAAAGAAATATTTATTTTAAATTAATTAAGAAAGAAATATGAGAAATTAAT 2917
QY 137 ----- 137
Db 2918 CAAGCAGAGAGATTTATCATATCTATGCGGATTAAGATATATATATGTAAAGAAA 2977
QY 137 ----- 137
Db 2978 CAGCACTACACATCTAATAATAAGTCTCATTAATGATATAAAAAATAGTGTAGCAAA 3037
QY 137 ----- 137
Db 3038 CAGTTATCCCTACAAAACCTTTTGGGGTAGATGATCCAGAGGTTGTTCCAGACTCT 3097
QY 137 ----- 137
Db 3098 TGCTTAATAAAGATTTTTTCTAATAATGATTGAAGAAATGTATATATGATGAATAAT 3157
QY 137 ----- 137
Db 3158 GAAAGAAAACATATCATATTTAAAAATATAAAGTATCAAGTAACGAATTAACAATA 3217
QY 137 ----- 137
Db 3218 GGAATATATCTATAAATGAATAATTTAGGCTTTTCGTTAATATCTTAGTTATTC 3277
QY 138 ----- 141
Db 3278 ATTTGTTCTTAATTTCCCTCTTATTTTTTTTGAATTAATCTATGACAGATTTTGGATC 3337
QY 141 ----- 161
Db 3338 CTTAAGAGCATTTAATAAATTTTCAGACTTACAAAGATGAGAGGCTTTGATGCTTTTGA 3397
QY 161 ----- 181
Db 3398 TAGCTTTCAAAATTTAGTTGTTGGAGAGAGAACTATCAATGCAATGCAAGATAG 3457
QY 181 ----- 192
Db 3458 GTGGCCAGATCTTCCAAATTAATAATCACTGTGTAATTTATTAACCTTGCTTAAGT 3517
QY 192 ----- 192
Db 3518 TTTAGCATATGTTGCTCGAATCTTTAACTGTTCTAAAGATATATATATTTGAAGA 3577
QY 192 ----- 192
Db 3578 GGTTGTCAAAATGATCATCATTTTGAAGATTCGCAACAAATATAGTTTATGTAATCTA 3637

QY 193 -----Pr 193
Db 3638 ATTTTCAGAGCATCTTCGCTTGTAAGTATGATCATGTTACCTTTTTTTCTTCATGAGCC 3697
QY 193 ----- 199
Db 3698 ATGCAGGATGACCAACGGTACGTTAATGTGATTTGATTAATAAAAAAGCCCTA 3757
QY 199 ----- 199
Db 3758 AAATATATTGAATTTTAATGAAAAGTAAATATATTTCTTAACCTTGGCAGAGACTAT 3817
QY 199 ----- 199
Db 3818 TACCCCTGCATTAATTAATAGTATTTTAAAGATATAAGTGTATAGTGAACAAA 3877
QY 199 ----- 199
Db 3878 AATTAGATTTCAAAAATTAATTTGAATAATTAATTAATTTTGCATATCA 3937
QY 199 ----- 199
Db 3938 ATATGATTAATAAATATATAGTTAAAGTCTTATGATTTGATCTTAATAAATAATCAATG 3997
QY 199 ----- 199
Db 3998 ACAAACAATAGTAGCGAGAAAGTATATAACAATACCTTTCAAGTAAATCATTTGT 4057
QY 199 ----- 199
Db 4058 ACACACACTCAAAACCTACGTTTTCTCGATTAATTTTCTATTTCTTAATAGTAA 4117
QY 199 ----- 199
Db 4118 TCAAGGCTATTAAGTCTGTCAAAATCTATACATTGAAACCTATCTTTGACGCTCGT 4177
QY 199 ----- 199
Db 4178 ACATTGAGATCGTTGAACAATGATGAATGATTTTAATCTTTGATTAATAATTA 4237
QY 200 ----- 212
Db 4238 ACTAATATTTGTTAATTTTCAGGCTTAACCTTCGAATTCAGAAAATTTGAAGTGA 4297
QY 212 ----- 232
Db 4298 TAACTTAAGATTAATAATGCAACAATTCATATCAAAATTTGAGTCATGCACTAATGT 4357
QY 232 ----- 252
Db 4358 TGTACCTTCAAAATTTGATGATCAATGCTTCAGCAAGAGCCCAATTAATGAGAGTCCA 4417
QY 252 ----- 268
Db 4418 TGTATCAATCTCAATATATTTCAATATCTGATATCTATTAATGGAACAGGTTATAT 4477
QY 268 ----- 268
Db 4478 TTAATTTTATTTATCAATTAATTAAGAAAAAGAGATATTTTATTTGATTAATCA 4537
QY 269 ----- 279
Db 4538 AATTATTAATTTTAATTTTTTTTATAGGTATATGATTAATTTGATTTCTGATC 4597
QY 279 ----- 296
Db 4598 TGAATATGACAGCCACAATATATCTTGTGTCAAGTATGCTATADAG-GTACTCTA 4656
QY 296 ----- 296
Db 4657 TTTTACAATATATCTGTTTCATTTCTATATTTCAATAAAGTATGATATATATAA 4716
QY 296 ----- 296


```

Db      4717 TTAAGTAAATCCTTAATTAATTATTTATGCAAAATTTTCTCTGTCTTTATGCTAAT 4776
QY      296 ----- 296
Db      4777 GACTAGACAAATTAATAGGCGCTTTGGATGGCGAATAAAGAGCTTTAAAAAGTA 4836
QY      296 ----- 296
Db      4837 CTTTAAAGTGTGAAACTTAATTTTAAATTAAGCAGTTATCGGTTTGATAAAGTGC 4896
QY      296 ----- 296
Db      4897 TGAAGTTTATGTCAAACGTGAAAAAGGAAAAATGAAAGAAAGATGTAGGTTATA 4956
QY      296 ----- 296
Db      4957 TGGGTTATTTGTATATAAAATATTAAGCACAAAAAGATAAATGTGTCAACTTAAACA 5016
QY      296 ----- 296
Db      5017 ACTTAAAGTACCTACCTACCTACCCAGCTTTAACTTTGGCTTAAATTAAGTTTAT 5076
QY      296 ----- 296
Db      5077 TTTTAAACTTAAATTAAGTTGTTTGAATTCGCAAGAGCTAATTAATGCAAAAAC 5136
QY      296 ----- 296
Db      5137 AGCTTTAAGTCAAGTTTGACAGGCTTTAAGCTGAGCAACAGGCTTTAAATGTCTG 5196
QY      296 ----- 296
Db      5197 CTATGATGTCTATATATTTAGCTTTTGAAGTAGTATATTATCTTAAGTTCA 5256
QY      296 ----- 296
Db      5257 CATTAATATACGTCTTAACATAGCACATATAGTTAATCAAAAGACAAATGATGATTA 5316
QY      296 ----- 296
Db      5317 TTTTGGCAATTTGATTAATTCACAAAGAAAGGATAGTTCAAGGTACATTCATGAT 5376
QY      296 ----- 296
Db      5377 TGAATATCATTAAGATCTAAATTTGAAGAAATCATTAATGAGGATCAAAATGTAT 5436
QY      296 ----- 296
Db      5437 TACTTATTAATACTATTCATTTTCATTTAATTAATTAAGAGTTTATA 5496
QY      296 ----- 296
Db      5497 ATCTAATAAATGATGCAATTAATTAATGACAAATGTGTTGTACTATTAATCTTCT 5556
QY      297 ----- 297
Db      5557 GAATATTTGCTATTTTCTCTTTTATTTTCCATGATTAATTAATGAGCTTGG 5616
QY      301 ----- 301
Db      5617 ATCTGAATTCAGAGCTTATGTCTTAATGTACTGTAATGAGCAAAATTAATCG 5676
QY      321 ----- 321
Db      5677 TGCCGAAATGAGTTAGATCAAGACTTGCA-GTACCTCCCCCCCCCCCC 5735
QY      332 ----- 332
Db      5736 CACAGGCCATTTTAAATTTTAAATTTTATTCGAATATCAATTAAGATTA 5795
QY      332 ----- 332

Db      5796 ATTGATTCATGTTGAAATTTATATTTGATTAAGATATGATTAATTAAGCTTTCAT 5855
QY      332 ----- 332
Db      5856 GTTATATAGAAAAAATGTTCAAGACTTCAGATTAATGCTACTAGTAAAT 5915
QY      332 ----- 332
Db      5916 GTCTGCTTTGTTAGAACTTTGTTATCCAGTTTGGTCATGATTAACCAACTTA 5975
QY      332 ----- 332
Db      5976 TAATGAAAGGGCTGCAACGCCGCCCACTAGTCTAGTATCAATAGAAAGTCTAC 6035
QY      332 ----- 332
Db      6036 GTCTGTTATTCAGATGACGCTTGTGTGAATTAATTAATTAATTAATTAACA 6095
QY      333 ----- 333
Db      6096 TGTATTAAGCATTAATTAATTAATGTTTAAATTAATGAGAGAGATCTGACAGC 6155
QY      338 ----- 338
Db      6156 TAGCAACATCAAAATTTCTGAATGTGAAATGCAAGAGCTTAATGATCCATTAATTA 6215
QY      358 ----- 358
Db      6216 CCAAACTATTTGATGAGTTGACCAATGATTAACAGGTATTTTATTAACGAC 6275
QY      370 ----- 370
Db      6276 AATTAATTAATTTTATTAATCTTAATCACTTACATCATTAATACTTGAGATTC 6335
QY      370 ----- 370
Db      6336 TTCACTAGTTAGTAATTTTGAATAGATTTTATGAATGATTAATTAATTAATTA 6395
QY      370 ----- 370
Db      6396 TTTTCTTAATTTATGATCTTTTGACATAGTCTAATAATCTTTTAAGTAACT 6455
QY      370 ----- 370
Db      6456 GAATATCATTAAGAAAAATGTTAGATTAATTAATTAATTTTATTAATTAAGCT 6515
QY      370 ----- 370
Db      6516 ATCTAATTTTGTATGTAAGATGGAACATATATCAATTAATTAATTAATTA 6575
QY      370 ----- 370
Db      6576 ATTAATCTTAATTAATTTGATTAATCAATCAATCAATCAATCAATCAATCAAT 6635
QY      370 ----- 370
Db      6636 CTCAAAATATATGATGATGATCAATTAACAAAGATTAATTAATTAATTAATTA 6695
QY      370 ----- 370
Db      6696 AAGAGAGATTTGGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 6755
QY      371 ----- 371
Db      6756 TCTCAGCATTAATTAATCACTTTGTCAGTTTGAAGCTTCAATTAATTAATTAATTA 6815
QY      381 ----- 381
Db      6816 TGTATGAATATCAAGGCAAGTGCACAAAGGTGCGCATTAATTTGATTCAGCA 6875
QY      401 ----- 401
Db      6876 CAACTTTCCATGTCAGAGATTTATTAATGGAATTAATTAATTAATTAATTAATTA 6935
QY      421 ----- 421

```


QY 361 TTYCAAPARVAlGluProCysIllegInPheSerAlaValGlnValIysAsnVal 380
Db 1120 TACTGTGACCAAGACAACTTGTCTCAGAACAGCTCGCTGTCCAGTGAATAATT 1179
QY 381 ValTYGluAsnIleIysGlyThrSerAlaThrIysValAlaIleIysPheAspCysSer 400
Db 1180 TCTTCAGAACATCAAGGACATGTGCTTCCAAATGCGCAATCACTTTGATTGACAG 1239
QY 401 ThrAsnProCysGluGlyIleIleMerGluAsnIleAsnLeuValGlyIleSerGly 420
Db 1240 AAGAGTTCCTCCATGCGAGGATGTGTGAAAGATGTGATCTTGAATGAGAGAGCT 1299
QY 421 LysProSerGluAlaThrCysIleAsnValIleIlePheAsnAlaGluIleValThrPro 440
Db 1300 GCAGAGCAAAAGCTTTGTGCAACATGTCAACTCACTGAAACAGAGATGTTTACCA 1359
QY 441 HisCysThrSerLeuGluIleSerGluAspGluAla 452
Db 1360 CATTGCCAGAA-----GAGAGAGAGAGAGCT 1386

RESULT 11
AY043233 1595 bp mRNA linear PLN 20-MAY-2003
LOCUS Vitis vinifera polygalacturonase mRNA, complete cds.
DEFINITION
VERSION AY043233
ACCESSION
KEYWORDS
SOURCE
ORGANISM
Vitis vinifera
Vitis vinifera
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; Vitaceae; Vitis.
1 (bases 1 to 1595)
Nunan,K.J., Davies,C., Robinson,S.P. and Fincher,G.B.
Expression patterns of cell wall-modifying enzymes during grape
berry development
Planta 214 (2), 257-264 (2001)
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
SUBMITTED (02-JUL-2001) Dept. Plant Science, University of
Adelaide, Waite Campus, Glen Osmond, SA 5064, Australia
3 (bases 1 to 1595)
Nunan,K.J., Davies,C. and Robinson,S.P.
Direct Submission
Submitted (02-JUL-2001) Plant Industry, CSIRO, Hartley Grove,
Adelaide, SA 5064, Australia
Location/Qualifiers
1. 1595
/organism="Vitis vinifera"
/mol_type="mRNA"
/cultivar="Shiraz"
/db_xref="taxon:29760"
55..1389
/codon_start=1
/product="polygalacturonase Pg1"
/protein_id="AAK81876.1"
/db_xref="GI:15081600"
/translation="WSLORRPVCLIVLSPFCYSLQASPPYDYLGEASGYDSQAN
HLHSTIOSRETPMPKRFSGKTSKSLASVKNVNVYGAKGSDSATERAFKAWKA
ACSPGSLVVPKKNVYLKPIITGPKSSITVOIYGVQASDRASNDMTNHLI
FENVONLAVOOGGTINGKMTWNSCKVNDLPCKGAPALATFYNKNLANVKNLKO
NAOQHWVFEKCVQASGLTITAPGNSPNTDGLHVSYTKIIOSSVIGTGDCTSI
VSGSKVQVNGITCGPGHISIGLSGSDSRAHSDVTNAGATISGTTNVRKTMWG
GSGASNIKKFONIMHNVENPIIIDOKYCCOSKCKGQSSAVOVONLVONIKGTSSS
KEATSLDSAKFPCQGIILRDIDIKGCGAATAVCSNARTVWGDVSPNCA"

ORIGIN
Alignment Scores:
Ptd. No.: 5.38e-92 Length: 1595

Score: 1204.50 Matches: 228
Percent Similarity: 75.74% Conservative: 53
Best Local Similarity: 61.46% Mismatch: 87
Query Match: 50.40% Indels: 3
Gaps: 3
US-10-691-374-2 (1-457) x AY043233 (1-1595)

QY 73 IleIysValIleAsnValLeuSerPheGlyAlaIysGlyIleAspGlyIleThrTyraAspAsn 92
Db 277 GTTAAATGTTAAAGTGAATTAATTATGAGCTAAAGGTGATGAAAGTAAAGCCACAGAG 336
QY 93 IleAlaPheGluGluAlaThrPheAsnGluIleIleCysSerSerAlaGlyThrProValGlnPheVal 112
Db 337 ---GCATTCAAGAAAGCTTGAAGGACAGCTGTTTCATCTCCAGAGATCTGTC---CTAGTG 390
QY 113 ValProIysAsnIleAsnValLeuLeuIleGluIleIleThrPheSerGlyProCysArgSer 132
Db 391 GTGCCCAAAACAAAGAACTATCTTAAAGCCAAATTCATTCAAGGATCTTGAATATCC 450
QY 133 SerIleSerValIlePheGlySerLeuGluIleSerIleIleSerAspTyr--- 151
Db 451 AGTATTACAGTGCAGATATATGAACTGTCAGGCACTCATCTGATCGCTCAAGT 510
QY 152 LysAspArgArgLeuTyrIleAlaPheAspSerValGlnAsnLeuValIleGlyIleGly 171
Db 511 AACGATATGACCCAGCTGCTTATCTTGAAGAAATGCCAAATTTAGCGCTTCAAGGTGT 570
QY 172 GlyThrIleAsnGlyIleAsnGlyIleValIleTyrProSerSerCysIleIleAsnIleSer 191
Db 571 GGCACATCAATGGAATAATGGAACAACGTGTGGAAACCTTGCACAAAGTGAATGAT 630
QY 192 LeuProCysArgAspAlaProThrAlaLeuThrPheThrAsnCysIleAsnLeuIleVal 211
Db 631 CTTCCTTGCAGAGGTGACCAAGCGCTTAACTTCTATACCTGCAAGACCTTGGCAGTG 690
QY 212 AsnAsnLeuIleSerIleAsnAlaGlnGlnIleHisIleIlePheGluSerCysThrAsn 231
Db 691 AAGAACTCAAGATCCAAATAATGCCAGCAATCATGTTCTTGAATAAATCGTGGGT 750
QY 232 ValValIleSerAsnLeuMetIleAsnAlaSerAlaIleSerProAsnThrAspGlyVal 251
Db 751 GTTCAAGCTTCGGGTCTCACCATTAAGTCCAGGGAACAGTCCCAACTGATGGAATC 810
QY 252 HisValSerAsnThrGlnTyrIleGlnIleSerAspThrIleIleGlyThrGlyAspAsp 271
Db 811 CATGTAGTTACACCAAAATTAATCCAGATTTCAAGCTCGATTATGGAACAGTATGAC 870
QY 272 CysIleSerIleValSerGlySerGlnAsnValGlnAlaThrAsnIleThrCysGlyPro 291
Db 871 TGTATTTCAATGTGTGATGATCCAGAGGTGCAAGTCAATGATCACTCGTGGACCA 930
QY 292 GlyIleGlyIleSerIleGlySerLeuGlyIleSerGlyIleAsnSerGluAlaTyrTyraSerAsn 311
Db 931 GGCCATGGAATCAGCATTTGGAAGCTTGAAGATTCAGAGCTTCATATTTTCGAGAT 990
QY 312 ValThrValAsnGluAlaIleIleIleGlyAlaGluAsnGlyValArgIleIleThrTyr 331
Db 991 GTTACAGTAAAGCGGCCACAGCTTTCGGGACCAAAATGAGGTGATGAATCAAGCATGG 1050
QY 332 GlnGlyIleSerGlyIleAlaSerAsnIleIlePheLeuAsnValGluMerGluAspVal 351
Db 1051 CAGGAGAGGTCTGGAAGCGCAAGCAACATCAATTTCAAGCACTTGAATCATCAATGATG 1110
QY 352 LysTyrProIleIleIleAspGlnAsnTyrCysAspArgValGluProCysIleGlnGln 371
Db 1111 GGAACCCCTAATAATATAGACCAAAAGTACTGTGACCAAGCAACATGCAATGACAG 1170
QY 372 PheSerAlaValGlnValIleAsnValValTyrGluAsnIleIleGlyThrSerAlaThr 391
Db 1171 AGCTCGGCTGTTCAAGTCAAAATGTGTGTAACCAATGAAGAAACAAAGTTCCTTCG 1230
QY 392 LysValAlaIleIlePheAspCysSerThrAsnThrProCysGluGlyIleIleMerGlu 411

DB 1231 AAGGAGCTATATCCGTTGATTGCAGTGGCAAGTTTCATGTCCAAAGCATTTCTGTCGA 1230
||||| ||||| |||||
OY 412 AsnleAenleuValGlyUserCilysPProSerGIuaIatThrCyAlAsnValHis 431
::|||.....
DB 1231 GACATTGATATTAAGAATTGGACGAGCGAAGAACGCCAAGCTGTATGCAACGATGCCAGA 1350

OY 432 PheAnaenAlaGlUnleValThrProHisCys 442
|||...|||...
DB 1351 GTGACTGTAAATGGGGGATGTGTCTCCAATATTC 1383

RESULT 12
AVOPOLYGAL 1725 bp mRNA linear PLN 09-AUG-1994
LOCUS Avocado polygalacturonase mRNA, complete cds.
DEFINITION L06094
ACCESSION L06094
VERSION GI:166950
KEYWORDS pectin depolymerase; pectinase; polygalacturonase.
SOURCE Persea americana (avocado)
ORGANISM Persea americana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophytes; Magnoliophyta; magnoliids; Laurales; Lauraceae;
Persea.
REFERENCE 1 (bases 1 to 1725)
AUTHORS Kutsunai,S.Y., Lin,A.C., Percival,F.W., Laties,G.G. and Christofferzen,R.E.
TITLE Ripening-related polygalacturonase cDNA from avocado
JOURNAL Plant Physiol. 103 (1), 289-290 (1993)
MEDLINE 94269193
PubMed 8208850
COMMENT Original source text: Persea americana (strain Haas) ripe fruit
meocarp cDNA to mRNA.
FEATURES
source location/Qualifiers
..1..1725
/organism="Persea americana"
/mol_type="mRNA"
/strain="Haas"
/db_xref="taxon:3435"
/ligase_type="mesocarp"
/dev_stage="ripe fruit"
1..33
34..1395
/EC_number="3.2.1.15"
/function="cell wall degradation"
/codon_start=1
/product="polygalacturonase"
/protein_id="AAA32914.1"
/translation="MALTRLPLPSILIMFCVSSHITLQKDPILLCVNGDPEFORAR
PTYPFPIIDERSISWGFPSSILSLERFNVPVGAPETSPDTDISVDFGARGGDTDRKA
FEKAKDCSSGSIVIPENKNYLKOTTFSGPKSDLRYIRIGRTIASPOSMVGH
NRKRIFEDISNLTBEGGTINGERTWMPSSCCRKSLEPCKAPRLATRSCKNLL
VSDLSIKDSQKHLSFDKCQDVASLNLTVAPEHSPTNDGIHTGTOKIHWSNVIGT
GDCCIISBSGRMVIATNTICGPBGHISIGSLGDRMSPAHSVGLVDEGNLPDTNGI
RIKTQGCGSGAKNIKPFONIIWNVNTNI IIDQYCDSKDPCPEQESAVKSNVAYNMMA
IRGTASEVAVKFDCKSPQGIILENNINLVNGCKETTMSCSNIVGTTEKVPPE
SCL"
1394..1725

ORIGIN
3'UTR
1394..1725

Alignment Scores:
Pred. No.: 1,36-87 Length: 1725
Score: 1153.00 Matches: 225
Percent Similarity: 72.35% Conservative: 68
Best Local Similarity: 55.56% Mismatches: 104
Query Match: 48.24% Indels: 8
DB: 8 Gaps: 5

US-10-691-374-2 (1-457) x AVOPOLYGAL (1-1725)

QY	48	PheGlnAlaIyTrLeuSerYTrLeuSerLyAsnIleGluSerAsnAsnAsnIleAspLyS	67
Db	208	ATGGAGATTGAGCCCTCCATCTTAGT-----TTGGAGCGGTTTAATCCAGTGGTGT	261
QY	68	ValAspLyAsnGlyIleLyValIleAsnValIleuSerPheGlyAlaLyGlyAspGly	87
Db	262	CCGGAACCTTACCAAGATACAGATATAGTGTGATGATTTTGGAGCCAGAGTCATGG	321
QY	88	LyAsThrYAspAsnIleAlaPheGluGlnAlaTrpAsnGlnAlaCysSerSerArgThr	107
Db	322	---ACAGATGACACAAAGGCATTGGAGAGGCTTCGAAGAGCGCTTGTTCATCAGACT	378
QY	108	ProValGlnPheValProLyAsnLyAsnYTrLeuLyGlnIleThrPheSer	127
Db	379	GTCTCC-----ATTGTCCTGAAAACAAGAACTATCTTTAAACAATTAACCTTTTCA	432
QY	128	GlyProCysArgSerSerIleSerValLyIlePheGlySerLeuGlnAlaSerSerLyS	147
Db	433	GGTCTTGTAATCTGATCTTGATCGTGTAAGATGGCGGAACGATTTGAAGCTTCAATG	492
QY	148	IleSerApLyTr-----LyAspArgArgLeuTrpIleAlaPheAspSerValGlnAsn	165
Db	493	CAATCAGATTGGGTGGGACATTAATCGAAMAAGATGGATTGAATTGAAGATCAGCAAC	552
QY	166	LeuValValGlyGlyGlyGlyThrIleAsnGlyAsnGlyIleValIlePTrProSerSer	185
Db	553	CTCACACTGAAAGCGGTGGAAACCATCATGGAATGGAGAGACATGTGGCATAGCTCC	612
QY	186	CysLySylleAsnLySerLeuProCysArgAspAlaProThrAlaLeuThrPheTrpAsn	205
Db	613	TGCAAAAGGAAGAAATCCCTTCCTTGGCAAAAGCCACCCTACGGGCTTAACTTCAGTCC	672
QY	206	CysLyAsnLeuLySylleAsnLeuLySerLySylleAsnIleGlnIleHisIleLyS	225
Db	673	TGCAGAAACCTGATTTGTAGGATCTCGATCAAGAGCACCACAAAGATCATCTTCT	732
QY	226	PheGluSerCysThrAsnValAlaSerAsnLeuNecIleAsnAlaSerAlaLySerS	245
Db	733	TTTGTAAAGTCCCAAGATCATAGCTTCTAATCTCATGATCGATGACACAGACATAGC	792
QY	246	ProAsnThrAspGlyValHisValSerAsnThrGlnTrIleGlnIleSerAspThrIle	265
Db	793	CCCAACACAGATGCGATTCATATACAGGCCACCAAAATTCATGTGATCAAAATTCCTG	852
QY	266	IleGlyThrArgLyAspAspCysIleSerIleValSerGlySerGlnAsnValGlnAlaThr	285
Db	853	ATCGGAACAGTGAATGATTTGATCTTCATTAAGAGTGAATCAAAATGCTATAGCCACA	912
QY	286	AsnIleThrCysGlyProGlyHisGlyIleSerIleGlySerLeuGlySerGlyAsnSer	305
Db	913	AACATTAACCTTGAGCCAGGCCATGAGATTAAGCATTTGGAGACCTTAAGATCAAAATTC	972
QY	306	GlnAlaIyTrValSerAsnValThrValAsnGlnAlaLySylleGlyAlaGlnAsnGly	325
Db	973	GAAAGTCAATGTTTACAGAGTACTTGTGGATGCGGCAACCTTTTCGATCAACAATAGGA	1033
QY	326	ValArgIleLySylleThrTrpGlnGlyGlySerGlyGlnAlaSerAsnIleLyPheLeuAsn	345
Db	1033	CTCAGGATCAAAACATGCGCAGGAGGTTCAAGAAAGTGCAAAAGATCAAAATTCCAAAAC	1093
QY	346	ValGluNecGlnAspValLySerTrpTrpIleIleIleAspGlnAsnYTrCysAspArgVal	365
Db	1093	ATTGTCAATGACCAACGCTCAAAATCTTAATATCAATAGTCAATATCTATGTGAATCCAA	1155
QY	366	GluProCysIleGlnGlnPheSerAlaValGlnAlaLySylleValIyTrGlyAsnIle	385
Db	1153	GACCATATCCGGAACAGATCAAGCTGTGAAGTAAAGCATATGTTGCTTACATGAACATC	1212
QY	386	LyGlyIyThrSerAlaTrnLyValAlaIleLySyllePheAspCysSerThrAsnPheProCys	405
Db	1213	CGTGAACAAATGCTCTGAGGTGCTGTAAATTCGATTTTCAGAGAGAGTCTTCCATGC	1272

Qy	406	Glucgyl1le1leWecGluhsenlleantleuVallgIyGusErG1yLyEProSeErGlu1ua	425
Db	1273	CAAGGATATTTGTGGAAAAACATCATCTGTTGGGAATGAGGAAAGGAACATCATG	1330
Qy	426	ThrCyLSyAsnVal	430
Db	1333	TCATCAGCAATATTT	1347
RESULT 13			
AY078936			
LOCUS	AY078936	1364 bp	mRNA linear PLN 09-MAR-2002
DEFINITION	Arabidopsis thaliana At2g41850/T11A7.5	mRNA, complete cds.	
ACCESSION	AY078936		
VERSION	AY078936.1	GI:19310404	
KEYWORDS	FLI CDNA.		
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
REFERENCE	Enkaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosid II; Brassicales; Brassicaceae; Arabidopsis.		
AUTHORS	Shim, P., Chen, H., Cheuk, R., Kim, C. J., Meyers, M. C., Banb, J., Bowser, L., Carninci, P., Chang, E., Dale, J. M., Goldsmith, A. D., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J. M., Lin, J., Miranda, M., Narusaka, T., Nguyen, M., Onodera, C. S., Palm, C. J., Quach, H. L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C. C., Toriumi, M., Wu, H. C., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R. W., Theologis, A. and Ecker, J. R.		
TITLE	Arabidopsis cDNA clones		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 1364)		
AUTHORS	Shim, P., Chen, H., Cheuk, R., Kim, C. J., Meyers, M. C., Banb, J., Bowser, L., Carninci, P., Chang, E., Dale, J. M., Goldsmith, A. D., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Lam, B., Lee, J. M., Lin, J., Miranda, M., Narusaka, T., Nguyen, M., Onodera, C. S., Palm, C. J., Quach, H. L., Sakurai, T., Satou, M., Seki, M., Southwick, A., Tang, C. C., Toriumi, M., Wu, H. C., Yamada, K., Yamamura, Y., Yu, G., Yu, S., Shinozaki, K., Davis, R. W., Theologis, A. and Ecker, J. R.		
TITLE	Direct Submission		
JOURNAL	Submitted (08-FEB-2002) Salk Institute Genomic Analysis Laboratory (SIGAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA		
FEATURES			
source	Location/Qualifiers		
	1..1364		
	/organism="Arabidopsis thaliana"		
	/mol_type="mRNA"		
	/db_xref="taxon:3702"		
	/chromosome="2"		
	/clone="At2g41850.6 (S500052)"		
	/ecotype="Columbia"		
	50..1351		
	/note="putative polygalacturonase"		
	/codon_start=1		
	/product="At2g41850/T11A7.5"		
	/protein_id="AA184942.1"		
	/db_xref="GI:19310405"		
	/translation="MARCTNLVFLMLLMLSMCKASRISPNVYDSYKRFKSDSLIKRREITIGRSFVARSLETPTVASDGAQDKDPTDTQA.FNAAWKAKSGNSAVNLLVPGKNTILKSLIOLGPGNSILTVQI.FGTLASAQKSDYDKISKTMFDSVNLSDGDDVUNGNETMWNQNSCKRNKAKPCTYATLTFPNSKSLIVKRLKVRNAQOIQI		
	STIEKSNVDSNVVATAPADSPNTQDHIHTNQNIRVESIIGTDDDISIESGQNIQ		
	QINDITCGEGHGISIGSLDDNSKSAFVSGVADGKLSGTQNGVAILKYQGSQNSN		
	ILFQNIQMDNVNPIIIDQDCDKCKTEKSAVQNVVYVADISGTASASNAITPNC		
	SKRYCCQGIIVLDRVNIKGGKATCTANVYVDKAVLPGQNSI"		
	104		
misc_difference	/note="compared to genome"		
	/replace="c"		
	445		
misc_difference	/note="compared to genome"		
	/replace="t"		
	445		
misc_difference	/note="compared to genome"		
	/replace="t"		

	misc_difference 1124	/note="compared to genome"
	misc_difference 1129	/replace="g"
	/note="compared to genome"	/replace="g"
Alignment Scores:		
Pred. No.:	2,136-87	Length: 1364
Score:	1149.00	Matches: 226
Percent Similarity:	67.71%	Conservative: 76
Best Local Similarity:	50.67%	Mismatches: 118
Query Match:	46.08%	Indels: 26
DB:	8	Gaps: 6
US-10-691-374-2 (1-457) x AY078936 (1-1364)		
QY	6 AasnSerIleLeuLeuLeuIleIleIlePheAlaSerSerIleSerThrCysArgSerAa	25
Db	65 AACCTGTACCGCTTTTCCTATTAGGGCTCTTTGATGCTCATTACAGTAAAGCTTCA	12
QY	26 ValIleAspAsnLeuPheLysGlnValTyr-----AspAsnIleu	40
Db	125 AAGATCAGCCCTTAAGTATGATGACATCTTATTAAGTTCAATCCATAGCTTAATC	18
QY	41 GtngIngluPheAlaIleAspPheGlnAlaTyrLeuSerTyrLeuSerIysAsnIleGlu	60
Db	185 AAGCAAA-----GAGACATACCGGAGTTGAGAGTTTGTAAGCCCTCTTGGCG	23
QY	61 SerAsnAsnAsnIleAspLysValAspLysAsnGlyIleLysValIleAsnValLeuSer	80
Db	239 ACTCCACCC-----ACGTTAGTGTCTGAT	26
QY	81 PheGlyAlaLysGlyAspGlyLysThrTyrAspAsnIleAlaPheGluGlnAlaTrpAsn	10
Db	266 TTGGAGCTTAAGAGATGAGAAACCGATGACACCGACGCGCTTCGTGAATCGTGAAG	32
QY	101 GlnAlaCysSerSerArgThrProValGlnPheValIleProLysAsnLysAsnTyrLeu	12
Db	326 AAGCATGTCTTCAATGAGCGGTAAATCTCTAGTTCCTTAAGCGAATCTTATCTC	38
QY	121 LeuLysGlnIleThrPheSerGlyProCysArgSerSerIleSerValIlePheGly	14
Db	386 CTAAAGCTTATCAATTAACTGGTGCATGCAATCTTATCTCACCGTTCAGATATCTCGT	44
QY	141 SerLeuGlnAlaSerSerLysIleSerArgTyrLysAspArgArgLeuTrpIleAlaPhe	16
Db	446 ACGTTATCCGATGTCATAAAGCATGCGATTACAAAGATATACGAAATGGATATGTTT	50
QY	161 AspSerValGlnAsnLeuValAlaGlyGlyIly-----GlyThrIleAsnGlyAsnGly	17
Db	506 GATGGCGTTACAATCATCATGATCGATGAGGCGCACATCGGGGTTGTGACGAAACGCC	56
QY	179 GlnValTrpTrpProSerSerCysLysIleAsnLysSerLeuProCysArgAspAlaPro	19
Db	566 GAAACGTGGTGGCAAACTCATGCAACCGAAAGCAAGGCTAACGCTTCGAAAGGCCCA	62
QY	199 ThrAlaLeuThrPheTrpAsnCysLysAsnLeuLysValAsnAsnLeuLysSerLysAsn	21
Db	626 ACGGCTCTTACTTTTAACTCAACCTGAAAGTCTGATGAGGAAGATCTGAAGTGAAGAT	68
QY	219 AlaGlnGlnIleIleIleLysPheGluLysCysTrpAsnValValAlaSerAsnLeuMet	23
Db	686 GCAACGAGATTCAGATTCGATTGAAAAGCTCAACGTTACGCTCTTAAGTGTGTG	74
QY	239 IleAsnAlaSerAlaLysSerProAsnThrAspGlyValIleValSerAsnThrGlnTyr	25
Db	746 GTAACTGCGCGTCCGATAGTCTTAAACCAACCATGATATTCATATCAACCAACCAAC	80
QY	259 IleGlnIleSerAspThrIleIleGlyThrGlyAspAspCysIleSerIleValSerGly	27
Db	806 ATTCCAGTCTCCGATCCATCATCTTGAACAGGCGATGATTCATATCTATTTGAAGTGA	85

QY 279 SerGlnAsnValGlnAlaThrAsnIleThrCysGlyProGlyValIleGlyIleSerIleGly 298
DB 866 TCACAAAAGTTCATCAATCAATATATATACCTTGGGTCGCGTCACCGTATCAATATGGG 925
QY 299 SerLeuGlySerGlyAsnSerGlnAlaThrValSerAsnValThrValAlaGlnAlaVal 318
DB 926 ACCCTTGGAGATGACATTCACAAAGCTTTGTCTCAGCGCTGACCTGTGATGTGTGAAG 985
QY 319 IleIleGlyAlaGlnAlaGlnValAlaGlyIleThrTrpGlnGlyGlySerGlyGlnAla 338
DB 966 CTTTCGCGTACACATGAGATGAGATCAAACTTACCAAGGAGGTCAGCAACTGCT 1045
QY 339 SerAsnIleValPheLeuAsnValGlnMetGlnAspValIleThrProIleIleIleAsp 358
DB 1046 ACCAATATATATATTCACAAAACATTCAGATGATTAAGATTCAGATCAATTAACGAC 1105
QY 359 GlnAsnTrpCysAspArgValGluProCysIleGlnGlnPheSerAlaValGlnVal 378
DB 1106 CAGACGACTCGACACAAAGGCAAA--TGCATCAACAGAAATCCGCGTCCAAAGTGAAG 1162
QY 379 AsnValValTrpGlnAlaGlnIleValSerAlaThrSerAlaThrValAlaIleValPheAsp 398
DB 1163 AAGCTGTGTACCGGACATTAAGTGCATCAGCGCATCGAAACGCAATTAAGTTAAC 1222
QY 399 CysSerThrAsnPheProCysGlyIleIleMetGlnAlaGlnIleAsnIleValGlyGlu 418
DB 1223 TCAGACAGAACTATCTCATGCGCAAGAAATGTGCTTGACAGAGTAACTTAAGGA-- 1279
QY 419 SerGlyValProSerGlnAlaThrCysValAsnValIlePheAsnAlaGlnIleVal 438
DB 1280 -----GGAAGACCACTTCACCAATGCTATGTGTATTAAGAGCTGT 1327
QY 439 ThrProHisCysThrSer 444
DB 1328 CTCCTCAGTGCACATCC 1345

RESULT 14
PAPOLYCU 1755 bp mRNA 11near PLN 11-MAY-1995
LOCUS P.americana mRNA for polygalacturonase.
DEFINITION X66426
ACCESSION X66426.1 GI:22630
VERSION
KEYWORDS polygalacturonase.
SOURCE Persea americana (avocado)
ORGANISM Persea americana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; magnoliids; Laurales; Lauraceae;
Persea.
1 (bases 1 to 1755)
Dopico, B., Lowe, A.L., Wilson, I.D., Merodio, C. and Grierson, D.
Cloning and characterization of avocado fruit mRNAs and their
expression during ripening and low-temperature storage
Plant Mol. Biol. 21 (3), 437-449 (1993)

JOURNAL MEDLINE 93184201
PUBMED 8095163
REFERENCE 2
AUTHORS Dopico, B.
TITLE Direct Submision
JOURNAL Submitted (27-MAY-1992) B. Dopico, Univ. of Nottingham, Dept. of
Phyiology & Envir. Sciences, School of Agriculture, Sutton
Bonington, Loughborough, Leicestershire, LE12 5RD, UK

FEATURES
source
1. 1755
/organism="Persea americana"
/mol_type="mRNA"
/cultiware="Hass"
/db_xref="taxon:3435"
/clone="SBA24"
/isue_type="pericarp"
/clone_lib="CDNA"
/dev_stage="ripening fruit"
49. 1437

CDS

/EC_number="3.2.1.15"
/codon_start=1
/product="polygalacturonase"
/protein_id="CAA47055.1"
/db_xref="GI:22631"
/db_xref="GOA:002096"
/db_xref="UniProt/Swiss-Prot:Q02096"
/translation="MALTRLLPILPISLPCFVSSHTTIOKPLICVNGDPGPDORAY
FTRPILIDERSLNGFSPSILSLERFPVPGSPSPDTDISVDPFARGDSDTRKA
FEKAWKDCSSGSVLIPEBNKYLLKQTFSGPCSKSLDKVIRGTHLEASDSDWVGH
NRKWIIFERDISNLTBGGTSSNGEITWDSCKRKSILPKCSAPLTFPSCKNL
VSDLSIKDSQKHLSPDKODVIALNLTAVTAPHEPNIDGHIITGTHVWNSVLT
GDGDISIESSKMVATNTCTGPGHISIGLSDNSSEAHVSGVLVDGNDLPTDGL
RIKMWGSGSAKNTIKKPTNIWNTNIIIDQYCDSPDEQSAVKVSNVAYM
IRGTSASVAKFPDCSKSPCOGYIVGININLVNGKKTTHSCSNIVGGLREGSLTF
LFMKRRVHECSY"

ORIGIN

Alignment Scores:
Pred. No.: 3,596-86 Length: 1755
Score: 1136.00 Matches: 223
Percent Similarity: 71.60% Conservative: 67
Best Local Similarity: 55.06% Mismatches: 107
Query Match: 47.53% Indels: 8
DB: 8 Gaps: 5

US-10-691-374-2 (1-457) x PAPOLYCU (1-1755)

QY 28 AspAspAsnIleuPheLysGlnValTrpAspAsnIleLeuGlnGlnIlePheAlaAsp 47
DB 166 GACCAAAAGGCTTACCACCATATTTGGTCCATTATGAT--GAGTTCTGACATC 222
QY 48 PheGlnAlaTrpLeuSerTrpLeuSerTrpLeuSerTrpLeuSerTrpLeuSerTrpLeuSer 67
DB 223 ATGGATTTGAGCCCTCATCTTGAGT-----TGGACCGGTTTAACTAGCGGTGT 276
QY 68 ValAspLysAsnGlyIleLysValIleAsnValLeuSerPheGlyValAlaGlyAspGly 87
DB 277 CCGAAACCTTCAACGATACATATTAATGATGATGATTTGGACCCAGAGTATGCG 336
QY 88 LysThrTrpAspAsnIleAlaPheGlnAlaTrpAsnGlnAlaCysSerSerArgThr 107
DB 337 ---ACAGATGACAAAGCATTTGAGAGCGTTGGAAGAGCGCTTTCATCAGATCT 393
QY 108 ProValGlnPheValAlaProLysAsnLysAsnTrpLeuLeuGlnIleThrPheSer 127
DB 394 GTCCCTC-----ATTGTGCTGAAAAAAGAACTATCTTAACTTAACCTTTTCA 447
QY 128 GlyProCysArgSerSerIleSerValLysIlePheGlySerLeuGlnAlaSerSerLys 147
DB 448 GGTCTCTGTAATCTGATCTTGTGTAAGATTCGCGAACCATTTGACCTTCATCTGAT 507
QY 148 IleSerAspTrp-----LysAspArgArgLeuTrpIleAlaPheAspSerValGlnAsn 165
DB 508 CAATCAGATGTTGGTGGACATATGCAAAACATGATGATTTGAACATCAGACAC 567
QY 166 LeuValValGlyGlyGlyGlyThrIleAsnGlyAsnGlyGlnValTrpTrpProSerSer 185
DB 568 CTCACATGAAAGGGGTGGACATCCATGAAATGAGAGACATGGGGATTAATCTCC 627
QY 186 CysLysIleAsnLysSerLeuProCysArgAspAlaProThrAlaLeuThrPheTrpAsn 205
DB 628 TCACAAAGAGAAATTCCTTCTTGCAAAACGACCTTACGGCTTAACTTCAGGTCC 687
QY 206 CysLysAsnIleuLysValAsnAsnLeuLysSerLysAsnAlaGlnIleHisIleLys 225
DB 688 TCACAAAGACCTATTTGTGAGCATCTCAGATCAAGAGACGCAAAAGATGATCTTCT 747
QY 226 PheGluSerCysThrAsnValAlaAspAsnLeuMetIleAsnAlaSerAlaLysSer 245
DB 748 TTTGATAGTGCACAGATGCTACGCTTATCTCATGATGATGACGACAGACATAC 807
QY 246 ProAsnThrAspGlyValIleValSerAsnThrGlnTrpIleGlnIleSerAspThrIle 265

Db	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	FEATURES																																															
Db	808	CCNAACACATGGCATTCAATTAACAGGCAACCAAGAAATCATCATGTGATGTAATTCGTGC	166	11eg1ThrGlyAspAspCysIleSerIleValSerGlySerGlnAsnValGlnIleThr	868	ATCGGAACAGGTGATGATTGTATCTCAATTAAGAAAGTGATCAAAAATGGTGTATGGCA	927	286	AsnIleThrCysGlyProGlyIleGlyIleSerIleGlySerIleGlySerGlyIleSer	928	AACTAATACCTGTGACACAGGCCCATGGAATTAAGCATTTGAAGCTTGAAGATCGAAATTC	987	306	GlnAlaIleValSerAsnValThrValAsnGlnAlaIleIleGlyAlaGlnIleGly	325	988	GAGGCTCATCTTCAAGAGTACTGTGGATGGCGGCAACCTTTTCATTAACAATATGA	1047	326	ValArgIleIleThrTrpGlnGlyIleGlySerGlyGlnAlaSerAsnIleIlePheLeu	1048	CTCAGATCAAAACAGGACAGGAGGCTTCAGGAAGTGCAAGACATCAAAATTC	1107	346	ValGlnMetGlnAspValIleThrProIleIleIleAsnGlnAsnThrCysAspArg	365	1108	ATTGTATGTCACACATCCAAATCTTAATCATTAATCAATACACTGTGATCTCC	1167	366	GluProCysIleGlnGlnPheSerAlaValGlnValIleAsnValValTrpGlnAsnIle	385	1168	GATCATATGTCGGAACAGAAATCAAGCTGTGAAGGTAAAGCATATGTTGCCATCA	1227	386	LysGluThrSerAlaThrTrpValAlaIleIlePheAspCysSerThrAsnPhePro	405	1228	CGTGAACAAAGGCGCTGAGAGGTGCTGTAATAATTCGATTGACGAAAGCTCTC	1287	406	GluGlyIleIleMetGlnAsnIleAsnLeuValGlyIleSerGlyIleProSerGln	425	1288	CAAGGATATATTGTTGAAACATCATCTGTTGGGATTCGAGGAAAGAAACACAT	1347	426	ThrIleIleAsnVal	430	1348	TCATGACGACATATT	1362	RESULT 15	BNA250919	1765 bp	mRNA	linear	PLN 31-AUG-2000
LOCUS	BNA250919	Brassica napus mRNA for endopolysaccharuronase (pegaz gene).	DEFINITION	AJ250919.1	GI:9967519	endopolysaccharuronase; pegaz gene.	Brassica napus (rape)	Brassica napus	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; euroside II; Brassicales; Brassicaceae; Brassica.	Gonzalez-Carranza, Z.H., Whitelaw, C.A. and Roberts, J.A.	Characterisation of a polysaccharuronase gene (PGA2BRAN) expressed during abscission in Brassica napus L., its relation to senescence and comparison with its homologous gene in Arabidopsis thaliana (PGA2AT)	unpublished	2 (bases 1 to 1765)	Gonzalez-Carranza, Z.H.	Direct Submission	Submitted (22-NOV-1999)	Gonzalez-Carranza Z.H., School of Biological Sciences, Plant Science Division, The University of Nottingham, Sutton Bonington Campus, PPS/P, Loughborough, Leicestershire LE12 5BP, United Kingdom	Location/Qualifiers	1. .1765	/organism="Brassica napus"	/mol_type="mRNA"	/culivar="Rafal"	/db_xref="taxon:3708"	/clone="CAM471"	/tissue_type="ethylene-promoted leaf abscission zone"																																		

```

gene
CDS
1..1765
/gene="pegaz"
73..1377
/gene="pegaz"
/codon_start=1
/product="endopolygalacturonase"
/protein_id="CAC05658.1"
/db_xref="GI:9967520"
/db_xref="GOA:O9FDY7"
/db_xref="UniProt/T-EMBL:O9FDY7"
/translation="MAHCTVTLGLYMSLFPMISWCEASRNNINRVDHSGYTRKSNL
IKRRDITRLKSVARASRLRPTTVSVDFEAKDGGKTDPTQAFVNMKRCSSGAVN
LLEBEGKITFLKSLRLNGPCKSVLTVOILDTLASOORSDYELSLKMTTFDGNSTLV
DGGATGVNNGAGFTWMWENSCRRNEAKCTKPAATLTFNRSKRLRVNNIRVADAOIQ
IIEKSCNVQSVNVEYVAPADSPNTGDIITNTQNIQVNSIIGTGDGDCISLGGSC
VOIIDLCPGHGHSIGSLADDDNSKA FVSGSVTYDNGAKLSTGDNGVR IKTNGGSGTAS
NIIFONTOMENVENPIIIDDDYCDKSKCTEOKSAVOIKRVVYNIISGTSASDIATFN
CSKNYPCQGIYLDKVINKGGKASCNSANWMDKGVLPCKCKST"

```

ALIGNMENT	ORIGIN	Alignment Scores:	Pred. No.:	1.16e-85	Length:	1765
Qy	11leuenuleuileilellephealasersetserthrcyarsersanvllie	Score:	1130.00	Matches:	221	
Db	103 tTACTAGTAGGCTCTTTTANAGCTCTAGTGCTGAGACCTTCAAGAAACACAT---	Percent Similarity:	67.34%	Conservative:	78	
Qy	28 Aspaspasnleuhelysglnvaltryaspasnilleuenglunglnphealanhisaep	Best Local Similarity:	49.77%	Mismatches:	117	
Db	160 -----AATGATATGATCTTCT	Query Match:	47.28%	Indels:	28	
Qy	48 Pheglmlatryleuser---tyrleuserlyasnilgeluserasnsanilleap	DB:	8	Gaps:	7	
Db	178 TATGGAACGTTCAAAATTCAGCTTAATCAAGCAGAGAGAGATATACCTGCTGAAA	US-10-691-374-2 (1-457) x BNA250919 (1-1765)				
Qy	67 Lyvalaslpasngilylleysval-----lleasnvalleuserphegilyla					
Db	238 AGCTTTCGAAGACCCTTTGGCGGCTTCCACACCGCTTAGTGTTCTGATTTGGAGCT					
Qy	84 Lyeglyaspdiylysthytryaspasnillealpheglunglnalatrpsanglnuacya					
Db	298 AAAGAGATGAGAAAACCGACGACGACGAGGCTTCGTGAATGATCGAAGAAAGACTGT					
Qy	104 SerSerArgthProvalglnphevalvalProlybasnlybasntryleuelysgln					
Db	358 TCTTCAAGTGGAGCTGTAAATCTCTTAGTCTCTGAAGGAAAACTATTTCTCAAGCT					
Qy	124 lilethPheSerGlyProcyarserserlileservalysillephegliserleu					
Db	418 CTTCGATTAATGGCCCATCGCAATCAAGTTCTTACGCTTCAGATTTAGAAACGTTACG					
Qy	144 AlaSerSerlylleserAsptrylybaspargargleutrpillealpheaspserval					
Db	478 GCTTCTCAACACGATCGGATTAATGAAGTCTCAGCAATGGATTACGTTGAGCGGCT					
Qy	164 Glnasnllevalvalglyglyly-----glythrilleasnglyasnely--glnval					
Db	538 AACAGCTTACCGTTGATGCGGAGCCACCGGAACTGTCAACGMAACGGGGCGGAAACG					
Qy	181 TrpTrpProSerSerCyllyslleasnlyserleuProcyarsargaspalaProthrpa					
Db	598 TGTGTGGAATCTCATCGAAACGACGACAGCTTAAGAAATCAAAAAGCCCTTAGCGCT					
Qy	201 leutnrphetrpsasnCyelyasnlleulyvalasnlleulysserlyasnalagln					
Db	658 CTTACTTTCVACCACTCGAAGATTTTAAAGATGATATATCTGAGCGTGAAGATCTGTA					


```
QY 221 GlnIleHisIleIyepheGluSerCysThrAsnValAlaSerAsnLeuMetIleAsn 240
Db 718 CAGATTTCAGATTTCGATTGAGAAATGCTCAACGTTCAAGTCTCAATATGTTAGAGTCACT 777
QY 241 AlaSerAlaIySerProAsnThrAspGlyValHisValSerAsnThrGlnTyrIleGln 260
Db 778 GCACCTGCTGATGATGCTTAACACTGATGATGATCCACATCATTACCTCAAAACATTCAA 837
QY 261 IleSerAspThrIleIleGlyThrGlyAspAspCysIleSerIleValSerGlySerGln 280
Db 838 GTCTCAATTCCATTCATTGGAACAGGTGATGATTGATATCTATTGAAAGTGATCACA 897
QY 281 AsnValGlnAlaThrAsnIleThrCysGlyProGlyHisGlyIleSerIleGlySerLeu 300
Db 898 AACGTTCAATCAACGATTAACTTACCTGGCCAGCCAGCGATATCGACATTGGAGCTTA 957
QY 301 GlySerGlyAsnSerGluAlaTyrValSerAsnValThrValAsnGluAlaIySileIle 320
Db 958 GGAAGATGACAAATTCCAAGGCAATTGATCAAGCGTGACTGTGATGATGCTTAACTTCC 1017
QY 321 GlyAlaGluAsnGlyValIArgIleIyThrTrpGlnGlyGlySerGlyGlnAlaSerAsn 340
Db 1018 GGTACAGACAAATGAGATTAGATCAAAACTTAACACAGGAGTTCTGGAATGCTGTAACAAC 1077
QY 341 IleIyepheLeuAsnValGluMetGlnAspValIySerTyrProIleIleIleAspGlnAsn 360
Db 1078 ATTATATTTTCAGAAACATTGAGATGAGATGTCGAGAAATCCAATCAATTAATTGACCMAAGAT 1137
QY 361 TyrCysAspArgValGluProCysIleGlnGlnPheSerAlaValGlnValIyAsnVal 380
Db 1138 TATTGCGACAAG---AGCAAAATGCACTGAACAAAATCAGCGGTCCAAATTAAGAACGTG 1194
QY 381 ValIyrgluAsnIleIyGlyThrSerAlaThrIyValAlaIleIyepheAspCysSer 400
Db 1195 GTGTACCGGACATTAAGCGGACACAGCGCATGCGACATAGCAATTAATTCAACTGCAAGC 1254
QY 401 ThrAsnPheProGluGluIyIleIleMetGluAsnIleAsnLeuValGlyGluSerGly 420
Db 1255 AADAACATATCCATGCGAAGAAATGTACTCGAACAAGTGACATTAAAGGA----- 1305
QY 421 LysProSerGluAlaThrCysIyAsnValHisPheAsnAsnAlaGluHisValThrPro 440
Db 1306 -----GGGAAAGCCAGTTGCAAGCAATGCTAATGTGATGATTAAGACCGGTGTTTACTT 1359
QY 441 HisCysThrSer 444
Db 1360 CAGTCAAAATCC 1371
```

Search completed: March 5, 2005, 00:17:23
Job time : 5020 secs

This Page Blank (uspto)

This Page Blank (uspto)